NEURAL BASIS OF PERCEPTUAL ORGANIZATION OF NATURAL SCENES: EMERGENCE OF OBJECT-BASED CODING IN THE PRIMATE VISUAL CORTEX

by

Jonathan R. Williford

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

November, 2014

Abstract

Discerning objects from their backgrounds is a fundamental process of vision. The coding of border-ownership in the early visual cortex is a neural correlate of this process. When stimulated with the contour of a figure, border-ownership selective neurons respond more strongly when the figure is on one side of their receptive field (the cell's "preferred" side) versus the opposite side of their receptive field (Williford and von der Heydt 2013). So far, border-ownership coding has only been shown with simple displays of geometric shapes (e.g., squares).

Here I studied border-ownership coding with static images of natural scenes by recording the responses of neurons in macaque visual cortex to occluding contours of objects in complex scenes. I found that subsets of neurons in visual area V2 signal the side of ownership of such contours. Decomposition of local and context influences in these neurons showed that the context-based border-ownership signals correlated with those for the edge of a square (which is locally ambiguous), but were weaker. I used stimuli with intermediate complexity along several dimensions to measure the relative influences of object shape, occlusion between objects, texture and color contrast to determine how they contribute to the border-ownership signal strength. I found that border-ownership signal decreases with the stimulus complexity. This

ABSTRACT

was especially pronounced when comparing a simple isolated square with a c-shape figure, overlapping squares, and natural stimuli. There were also smaller decreases when changing from uniform squares to natural texture squares and from squares to silhouettes of natural shapes. In conclusion, subsets of neurons in V2 do code for the border-ownership in natural scenes, however, the strength and accuracy of these early estimates of border-ownership decreases with the complexity of the visual stimulus.

Thesis Advisor: Rüdiger von der Heydt

Thesis Reader: Ernst Niebur

Acknowledgments

The time I have spent as a PhD student has been one of the best times of my life, in no small part because of my wife and son, the people in my lab, and the Johns Hopkins community at large. We are at a very interesting point in history, where scientific discovery is progressing at an astounding rate. I have been fortunate to be able to contribute a small part to this scientific progress and to pursue the research questions that interest me.

We stand on the shoulder of giants. Life as we know it would not be possible without the scientists and others that have contributed to our knowledge throughout history, some of whom have even been persecuted for advancing scientific knowledge in the darker passages of human history. I would like to thank all of those who have contributed to science.

Much of my interest in neuroscience has been driven by my experience in computer vision, which emphasized the difficulty of the problem of visual perception. I became particularly interested in how our visual system was able to organize the many low-level features that occur in the visual scene into objects. Dr. Rüdiger von der Heydt's research interests were the perfect match for me to explore this level of the visual system, and I am very fortunate that I was able to join his lab.

Rüdiger has been a great mentor and friend. I have been very lucky to have him as

my mentor, not only because of the overlap in our research interests, but also because of his kindness, encouragement, and constructive criticism. Rüdiger's guidance and insights have been indispensable. While he often allowed me to find solutions to my own problems, I knew he was always there to offer support when I wanted it. His knowledge of the visual system is incredible, and he has been a wonderful resource. He instilled in me an appreciation of the early Gestalt psychologists, who can serve as a guide to visual neuroscience research to this day. He will continue to serve as a role model throughout my career.

My fellow lab members have also always been around to offer help. Philip O'Herron and Anne Baldwin Martin, as well as Rüdiger von der Heydt, have taught me how to train monkeys and record neurons with extracellular microelectrodes and, with the help of Fangtu Qiu, have offered countless technical support in debugging problems with the recording. Fangtu wrote the experiment control and recording system, VLAB. Hee-kyoung Ko helped with setting up the experiment and performed some of the recording. Ofelia Geralde, the animal research specialist in our lab, offered a lot of assistance in preparing monkeys for recording and taking care of the monkeys. She, amazingly, also created the microelectrodes used in the experiment, which are the envy of the Mind/Brain Institute. I would also like to thank Hao Lei for technical support with the hardware.

I would also like to thank my thesis committee members: Marshall G. Hussain Shuler (chair), Ernst Niebur (thesis reader), Stewart H. C. Hendry, and Howard

ACKNOWLEDGMENTS

Egeth. They offered many insights and helpful critiques and suggestions. They have been very supportive, which has encouraged me to continue to pursue my career goals in academia.

I would like to thank my mom and dad for their love and support. My dad taught me the value of hard work, which has served me well throughout my life. I am very fortunate that both of my parents have always been there for me.

I have also benefitted greatly from many open source projects. This thesis has been written using LaTeX and knitr (Xie 2013, 2014), which allows for the analysis results and plots to be updated automatically through its many versions.

I am also profoundly grateful for my wife, Michelle. She has been very supportive of me and my career goals. Both she and our son have made my life much fuller and more enjoyable. I am very proud of my son. He has already been presented with many challenges. He has amazed me with his ability to rise up to these challenges even at his young age.

Dedication

This thesis is dedicated to my wife, Michelle, and our son.

Contents

Abstract **ii**

Acknowledgments **iv**

List of Tables **xii**

List of Figures **xiii**

1 Introduction **1**

1.1 Border-ownership cues 8

1.2 Border-ownership coding 11

1.2.1 Configural cues 11

1.2.2 Stereoscopic cues 17

1.3 Persistence and remapping of border-ownership signals 19

1.4 Neural models and constraints 21

1.4.1 Feedforward models 21

1.4.2 Lateral propagation models 24

1.4.3 Feedback models 24

1.5 Summary of thesis 26

CONTENTS

2 Methods **28**

2.1 Preparation 29

2.2 Recording setup 30

2.3 Eye tracking 31

2.4 Presentation of stimuli 31

2.5 Behavioral paradigm 32

2.6 Mapping of the classical receptive fields 32

2.6.1 Orientation mapping 33

2.6.2 Position mapping 34

2.6.3 Color mapping 34

2.7 Side-of-object selectivity tests 35

2.7.1 Classical side-of-object tests using geometric shapes 36

2.7.2 Side-of-object selectivity tests using natural scenes 38

2.7.2.1 Color space manipulation for local contrast 43

2.7.3 Side-of-object selectivity tests using "hybrid" geometric-natural stimuli 45

2.7.3.1 Texture synthesis 46

2.8 Data analysis 48

2.8.1 Analysis of variance 49

2.8.2 Analysis of neural population 51

2.8.2.1 Time course analysis 52

2.8.3 Distribution of side-of-object effects across scene points 52

2.8.3.1 Within a single cell 53

2.8.3.2 Over cell population 55

3 Border-ownership coding of natural scenes in primate V2 57

3.1 Introduction 57

3.2 Side-of-object selective cells 60

3.2.1 Example cells 63

3.3 Border-ownership coding: consistency of side-of-object with standard test 64

3.4 Time course 68

3.5 Side-of-object responses across scene points 68

3.6 Surround suppression uncorrelated with side-of-object modulation . . 73

3.7 Effect of image rotation on neural responses 75

3.8 Conclusion 77

4 Border-ownership coding of geometric, natural, and "hybrid" stimuli in primate V2 78

4.1 Example cell 79

4.2 Stimulus comparisons across neural population 82

4.2.1 Responses to geometric shapes 82

4.2.2 Effect of natural color and texture on squares 84

4.2.3 Effect of uniform textures and silhouettes 89

4.2.4 Comparison of silhouettes with natural scenes 90

4.3 Conclusion 91

5 Conclusion **93**

5.1 Summary of findings 93

5.2 Discussion 96

5.2.1 Contribution to understanding border-ownership coding and perceptual grouping 97

5.2.2 Contribution to understanding visual processing of natural images 99

A Border-ownership coding of natural scenes in primate V1 **103**

B Comparison across monkeys **105**

Glossary **115**

Bibliography **117**

Vita **126**

List of Tables

2.1 The number of V1 and V2 neurons presented with both the standard square and natural scene side-of-object selectivity tests, by monkey and hemisphere. 29

List of Figures

1.1 Example natural scene with visual occlusions. 2
1.2 Bregman's inkblotted Bs 3
1.3 Rubin face-vase illusion 4
1.4 Hypothetical post-stimulus histograms (PSTHs) of orientation selective neuron. 6
1.5 Border-ownership cues: convexity, area, symmetry, enclosure 10
1.6 Border-ownership cues: wide base, lower region, protrusion 11
1.7 Border-ownership cues: spatial frequency, extremal edges 12
1.8 Illustration of stimuli from the geometric tests. 13
1.9 Summary of the results from Zhou, Friedman, and von der Heydt (2000) using the standard (single square) test. 15
1.10 Transparent overlay stimuli. 16
1.11 Short-term memory of border-ownership. 20
1.12 Diagrams of three border-ownership models. 22

2.1 Images with scene points used for natural and natural-geometric hybrid experiments. 40
2.2 Manipulations of an example natural scene point. 41
2.3 Profile of the image patches. 43
2.4 Example hybrid versions of tiger scene. 45
2.5 Analysis of the side-of-object distribution among scene points for simulated cell. 53

3.1 Responses from cell 27lj2d to standard square and natural scenes. . . 61
3.2 Responses from cell 28ik1a to standard square and natural scenes. . . 62
3.3 Population comparison of side-of-object effects between standard square and natural stimuli. 65
3.4 Population comparison of border-ownership *modulation indices* between standard square and natural stimuli. 67
3.5 Time course of responses elicited by standard squares and natural scenes (by global context). 69
3.6 Scene points sorted by side-of-object response difference for cell 27lj2d. 71
3.7 Scene points sorted by side-of-object response difference for cell 28ik1a. 72
3.8 Scene points sorted by side-of-object response difference for population of V2 border-ownership cells. 73

LIST OF FIGURES

3.9 Comparison of surround suppression and the side-of-object modulation produced by the global context of natural scenes. 74
3.10 Effect of image orientation on neural responses 76

4.1 Responses from cell 28ik1a to geometric stimuli. 80
4.2 Responses from cell 28ik1a to natural-geometric hybrid stimuli. . . . 81
4.3 Population comparison of border-ownership effects between different geometric stimuli 83
4.4 Population comparison of border-ownership *modulation indices* between geometric stimuli 85
4.5 Population comparison of border-ownership effects between standard square and hybrid stimuli. 86
4.6 Population comparison of border-ownership *modulation indices* between standard square and hybrid stimuli. 87
4.7 Summary of comparison of border-ownership effects elicited by different stimuli. 88
4.8 Population comparison of side-of-object effects when adding texture and silhouette shape. 89
4.9 Population comparison of side-of-object effects of silhouette shapes and natural scenes. 90

A.1 V1 population comparison of side-of-object effects between standard square and natural stimuli. 104

B.1 Monkey M27: Comparison of side-of-object effects between standard square and natural stimuli. 106
B.2 Monkey M28: Comparison of side-of-object effects between standard square and natural stimuli. 107
B.3 Monkey M29: Comparison of side-of-object effects between standard square and natural stimuli. 108
B.4 Monkey M27: Population comparison of border-ownership effects between different geometric stimuli 109
B.5 Monkey M28: Population comparison of border-ownership effects between different geometric stimuli 110
B.6 Monkey M29: Population comparison of border-ownership effects between different geometric stimuli 111
B.7 Monkey 28: Population comparison of border-ownership effects between standard square and hybrid stimuli. 112
B.8 Monkey 29: Population comparison of border-ownership effects between standard square and hybrid stimuli. 113
B.9 Summary of comparison of border-ownership effects elicited by different stimuli across monkeys 114

Chapter 1

Introduction

Our brains are so amazing in their ability to process visual scenes that it is hard not to underestimate the difficulty of visual perception. Our eyes and retinas are incredible in their ability to detect and process visual information over the huge range of luminance that our eyes are able to see: from the ability to perceive a single photon of light when in complete dark (Rieke and Baylor 1998) to being able to make sense of the bombardment of photons in bright daylight. This, however, is only the very beginning of visual perception. Perhaps the most fundamental problem of visual processing is organizing the low-level primitive image features from our retina into objects.

During the projection of light to our retinas, coherent objects, such as the panda cub in Figure 1.1, can occupy multiple separate visual regions on the retina. The brain's process of assigning borders to objects and inferring the original visual structure of a scene is called perceptual organization, and occurs automatically and with limited conscious influence.

Resolving which objects own the visual borders in a scene is necessary for recog-

Figure 1.1: Example natural scene with visual occlusions. Perceptual organization allows us to effortlessly perceive the panda cub as a coherent object, even though there are many separate visual regions that make up the cub (Balleis 2013).

nizing shapes. Figures 1.2 and 1.3 demonstrate this importance. Figure 1.2 shows the same blue regions in A and B. In Figure 1.2A, each blue region appears to own its entire border (and is hence perceived as a separate object), while in Figure 1.2B, the black ink blot takes ownership of the borders. This tells the visual system that the borders along the black ink blot do not define the blue shapes, allowing the perception of letters behind the ink blot. Even after you learn the identity of the blue letters, it is difficult to see the letters in Figure 1.2A.

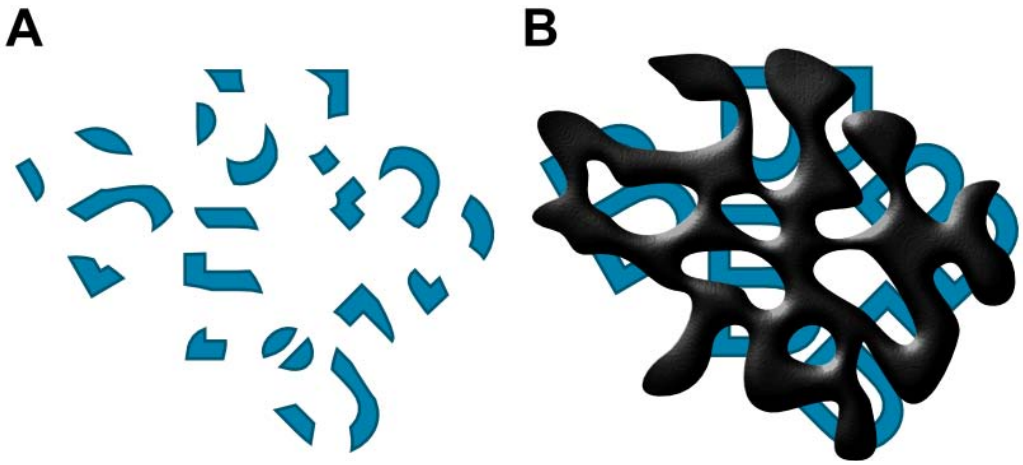


Figure 1.2: Bregman's inkblotted Bs (Bregman 1981). The blue regions are exactly the same in A and B. However, the black ink blot takes ownership of blue-black borders, making the blue regions appear to continue underneath and allowing the visual system to perceptually group these regions together.

In Figure 1.3, you can either see two black face profiles or a single white vase. In this case, the border-ownership is ambiguous. When the borders are perceived as being owned by the faces, you perceive the shape of the faces but not of the shape of

the vase. When the borders are perceived as being owned by the vase, you perceive the shape of the vase but not of the shape of the faces.

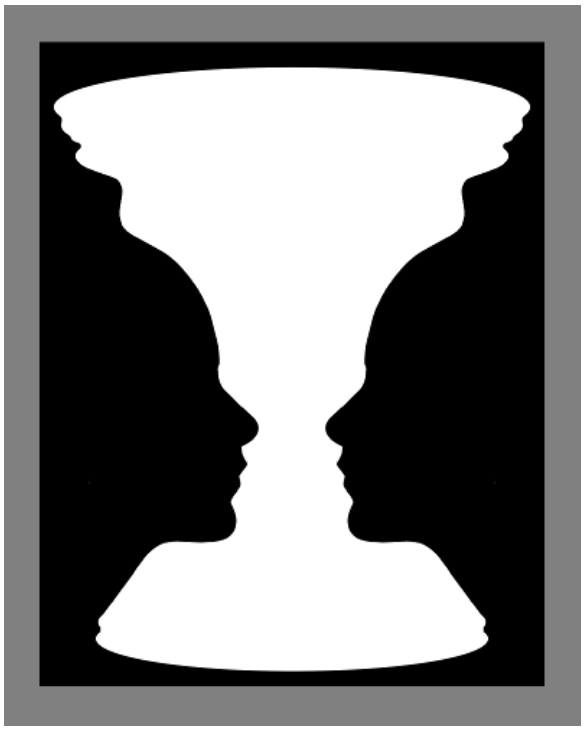


Figure 1.3: Rubin face-vase illusion (Rubin 1915). The border-ownership is ambiguous such that you can perceive either two black face profiles or a single white vase.

Knowing how our brains are able to accomplish this perceptual organization will undoubtedly give helpful insights for the computer vision community. For example, it will help that community solve difficult tasks such as pedestrian and object detection and tracking, without relying on expensive and limited "active" sensors, such as LADAR (laser radar), and could lead to smarter, life-saving automotive collision avoidance systems and rescue robots. In addition, this line of research may help us to

understand neurological visual impairment, also known as cortical visual impairment, (Good et al. 2001), and lead to the development of therapies.

Fortunately, there are two key phenomena discovered in the past 20 years that give insights into the neural basis of perceptual organization: border-ownership coding and figure-ground modulation. This thesis research focuses on border-ownership coding, which appears earlier in time and is not dependent on attention. It was discovered by Zhou, Friedman, and von der Heydt (2000) and was found to occur in orientation selective (Figure 1.4A) neurons in areas V1, V2, and V4 of primates. The responses of each of these neurons are driven by a small region of the visual field called the neuron's classical receptive field (CRF), so named because it was originally believed that the responses were only affected by the stimuli within the CRF. We now know that the surrounding context can significantly modulate the neural responses, such as by border-ownership. The CRF is now defined as the region in the visual field that drives a neuron's response.

Zhou, Friedman, and von der Heydt (2000) found that when a border of a figure is presented to the CRF of a neuron, that neuron will fire at a higher rate when the figure is on one side of the CRF compared to the opposite side (Figure 1.4B). The side where the figure elicits a higher firing rate is called the preferred side-of-figure. For example, the cell in Figure 1.4B "prefers" the figure to be on the upper right, as determined by its higher firing rate on this side compared to the lower left. This preference is fixed, much like their preferred orientation (Figure 1.4A).

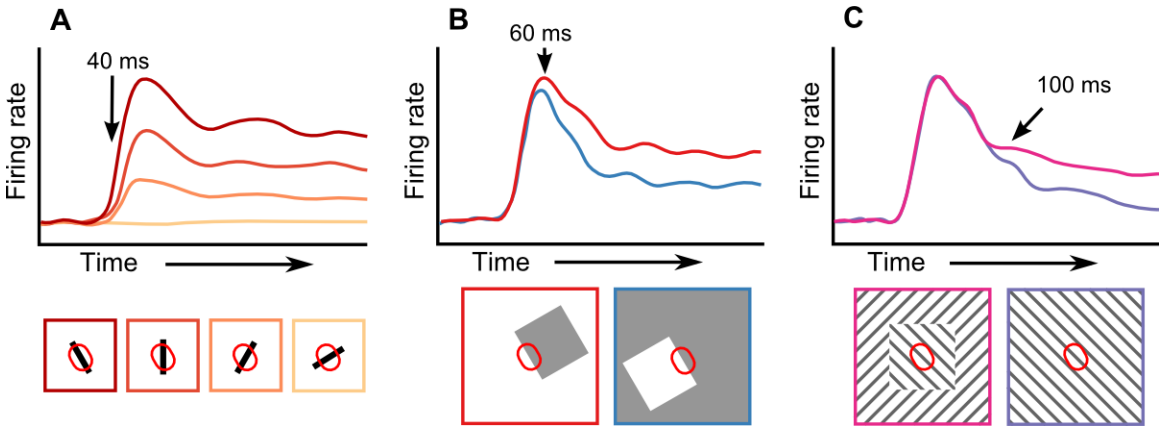


Figure 1.4: The post-stimulus histograms (PSTHs) of a hypothetical orientation selective neuron, which plot the neuron's firing rate in time after the presentation of a static stimulus. The displays are shown below the PSTHs, where the color of the enclosing boxes matches the corresponding response curve. (A) A response showing orientation selectivity. The difference in firing rate is apparent as soon as the neuron starts responding to the stimulus, which is around 40 ms after the stimulus onset for rhesus macaques. (B) A response showing side-of-object selective (also called border-ownership selectivity). The border-ownership signal, which is the difference in the firing rates of when the figure is on one side versus the other, begins to appear around 50 ms. (C) A response showing a related phenomena, called figure-ground modulation (V. A. Lamme 1995). The figure-ground modulation signal appears later, around 80 ms near the borders and 100 ms in the center of the figure.

So far, border-ownership coding has only been shown with simple displays of geometric shapes (e.g., squares). In this thesis work, I studied the border-ownership coding of natural scenes in visual area V2, using microelectrodes to record from isolated neurons. The visual system has evolved to process information from the real world. Obviously then, if a theory of the visual system cannot generalize to natural scenes and can only explain phenomenon using artificial stimuli in artificial settings, then this theory will have limited applicability.

This thesis establishes that the visual area V2 of primates does indeed code for border-ownership in natural scenes, and that the global context is used for this assignment. It also shows that the strength of the border-ownership signal varies by the stimulus type. An isolated square is shown to elicit the strongest signal. More complex geometric stimuli, such as the c-shape figure and overlapping squares elicit a significantly weaker signal across the population. Similarly, the global context of natural scenes elicit a significantly weaker signal. I found that local cues did not seem to contribute to the border-ownership coding across the population.

The rest of this introduction will review some of the findings and models of border-ownership coding in the early visual cortex and then summarize the remaining chapters.

1.1 Border-ownership cues

Psychologists have discovered many cues that our visual systems uses to assign border-ownership, often called figure-ground perception in this literature (see Peterson and Salvagio 2010 for review). These cues become especially pertinent when border-ownership is studied in natural images.

Some of the local cues include:

1. Convexity cue - concave side of a curve is more likely to be seen as the figure (Figure 1.5a).

2. Extremal edge / self-occluding cues (Palmer and Ghose 2008) - as the convex surface of an object wraps behind itself, out of the view, it often causes textures and shadows to follow the border of the object (Figure 1.7b).

3. Lower region cue (Vecera, Vogel, and Woodman 2002) - there is a bias for the lower region of a horizontal border to be perceived as the figure (Figure 1.6b).

4. T-Junction cues (see Albright and Stoner 2002 for review) - end-stopped or T-junction type lines indicate that an edge is being occluded, hence part of the background.

5. Color cues (Troscianko et al. 1991) - saturation influences the perception of depth.

6. Protrusion cue - region with protrusions tend to be seen as the figure (Figure 1.6c).

7. Spatial frequency and relative depth-from-focus cues - Klymenko and Weisstein (1986) showed that higher frequency regions tend to be perceived as occluding region, at least for sharp (high frequency) borders (Figure 1.7a). Marshall et al. (1996) showed that when the border is blurry, that the region that is similarly blurry tends to be seen as the occluding region. This suggests that blur due to the relative depth from the focus can be used to as an occlusion cue.

There are also the following global cues:

1. Contrast cues (O'Shea, Blackburn, and Ono 1994) - the region that has the highest contrast with the background will tend to be seen the figure.

2. Areas of the regions - smaller regions tend to be perceived as the figure (Figure 1.5b).

3. Symmetry cues - symmetric borders tend to make the region in-between be perceived as the figure (Figure 1.5c).

4. Enclosure cue - enclosed regions tend to be perceived as the figure (Figure 1.5d).

5. Large base cue - region with larger base tends to be seen as the figure (Figure 1.6a).

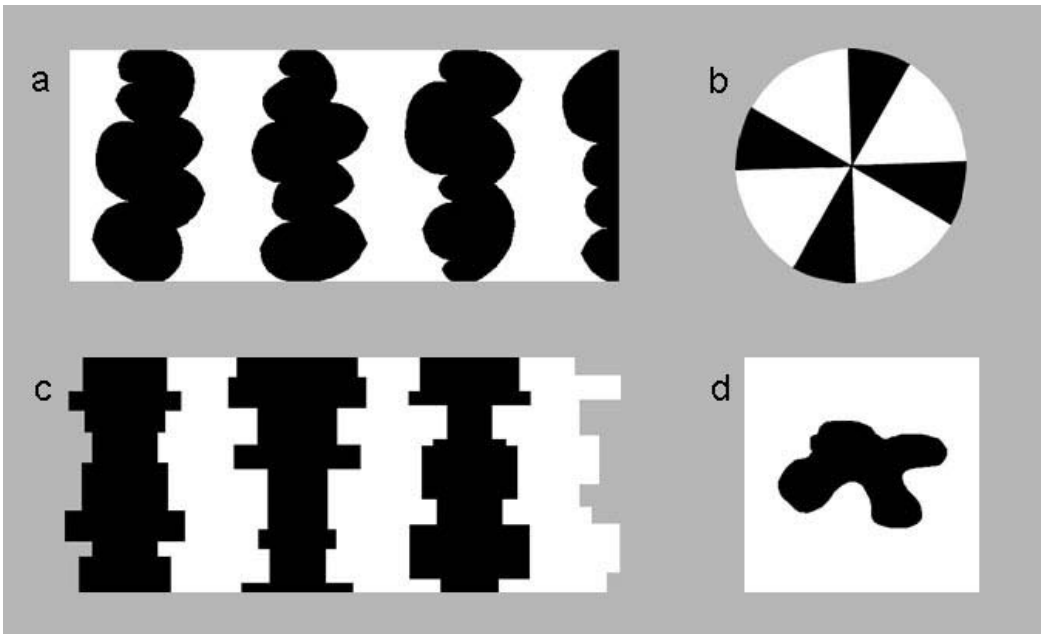


Figure 1.5: Border-ownership cues. We are more likely to see a region as owning the border when the border is concave towards the region, when the region is smaller, when there is symmetry with another border, and when the region is enclosed by the border. Reprinted from (Peterson and Salvagio 2010) with permission.

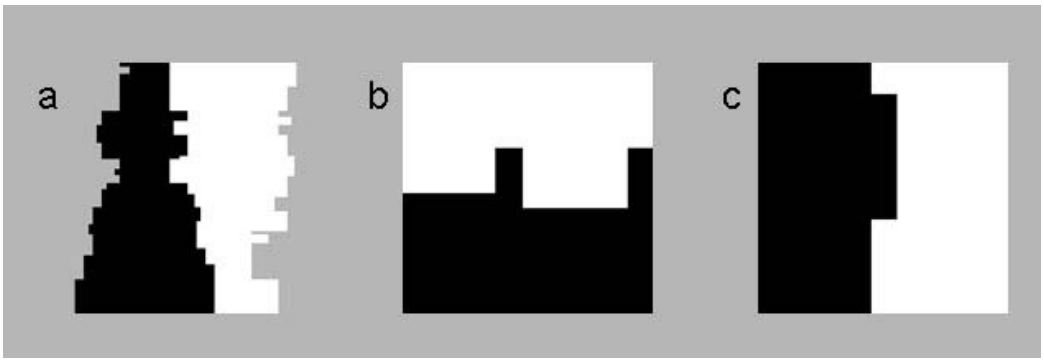


Figure 1.6: Border-ownership cues. a) Regions that have a larger base are more likely to be seen as owning the border. b) If two regions form a horizontal border, the lower region is more likely to be perceived as owning the border. c) Regions with protrusions are more likely to be seen as owning the border. Reprinted from (Peterson and Salvagio 2010) with permission.

1.2 Border-ownership coding

1.2.1 Configural cues

Zhou, Friedman, and von der Heydt (2000) recorded isolated single cell neuronal activity from awake behaving macaques. For the initial experiment, called the standard test, the border of a uniform colored square was aligned to the receptive field of the neuron, rotated to match the neuron's preferred orientation. The stimulus within the CRF was kept constant while changing the side of the figure by concurrently reversing the colors of figure and ground, as shown in Figure 1.8 (compare top and bottom dis-

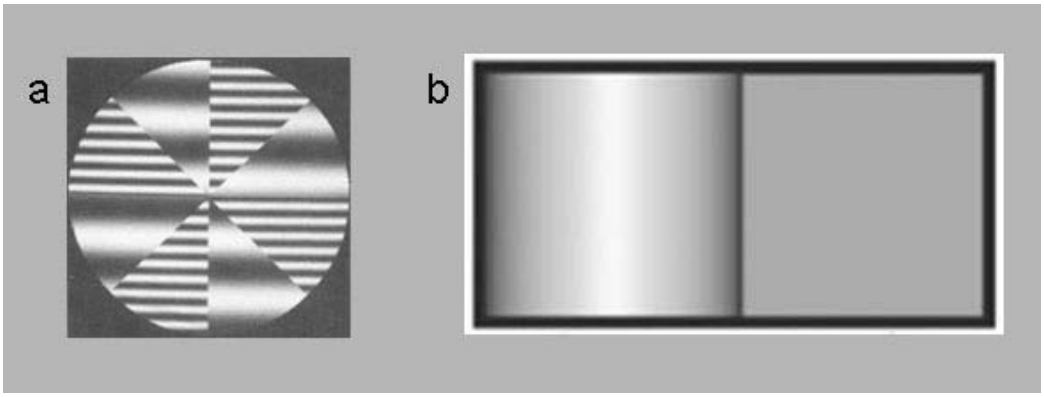


Figure 1.7: Border-ownership cues. a) Regions with higher spatial frequency are more likely to be seen as owning the border (Klymenko and Weisstein 1986, however, see Marshall et al. 1996). b) Regions with extremal edges are more likely to be seen as owning the border (Palmer and Ghose 2008). Reprinted with permission from (Peterson and Salvagio 2010) and respective journals.

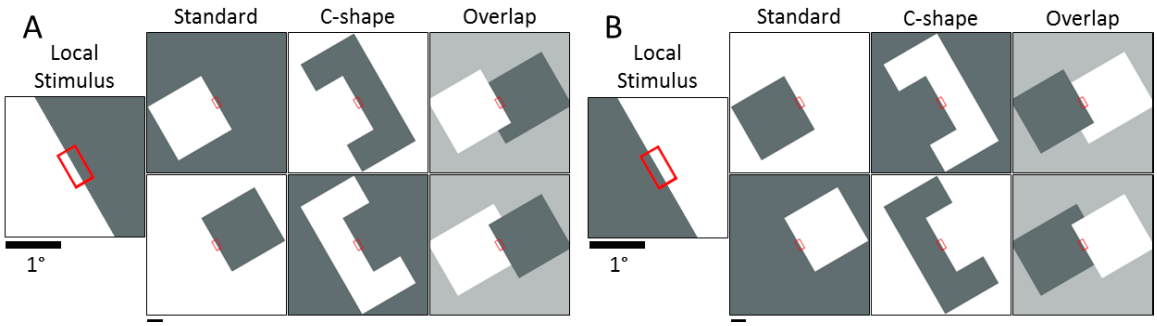


Figure 1.8: Illustration of the stimuli for a neuron that prefers borders oriented at 120°. The mean receptive field size of V2 neurons recorded in (Zhou, Friedman, and von der Heydt 2000) (.7° × .4°) is denoted by the red rectangle (not shown to subject).

All stimuli within each sub-figure, (A) and (B), are identical in the region surrounding the CRF, as shown in zoomed in display on the left of each sub-figure. The single square stimuli were used as the standard test of border-ownership coding, while the C-shape and overlapping square stimuli were shown to some of the neurons for comparison.

plays). The border-ownership signal of a neuron was defined by the response to locally identical stimuli when the figure is on its preferred side compared to non-preferred side. For example, in the neuron shown in Figure 1.4B, the border-ownership signal is the difference between the red and blue line. Zhou et al. studied the visual areas V1, V2, and V4 and found border-ownership selective neurons in each area (Figure 1.9). In V2, these were more than 50% of the orientation selective neurons that responded to contrast edges (which constitute about 80% of all V2 neurons). In V1, less than 20% were border ownership selective. In V4, the fraction was around 50%, but this is the percentage of neurons that could be activated with figure edges, which were about half of the cells encountered.

Due to the large proportion of border-ownership selective neurons and the amount of receptive field overlap in the early visual areas, there will be many neurons whose receptive field encode a given piece of a figure's border. Roughly half of these neurons will prefer the figure to be on one side of the border, while the other neurons will prefer the opposite side. The actual side of the figure will then be encoded at each location by the difference of the firing rates of the neurons in two pools with opposing side preferences.

In addition to the standard test, some cells were also presented with C-figures and overlapping squares, like those shown in Figure 1.8. Both of these displays elicited a significant border-ownership in a smaller proportion of cells than the single square. However, when both the single square and one of the other figures elicited significant

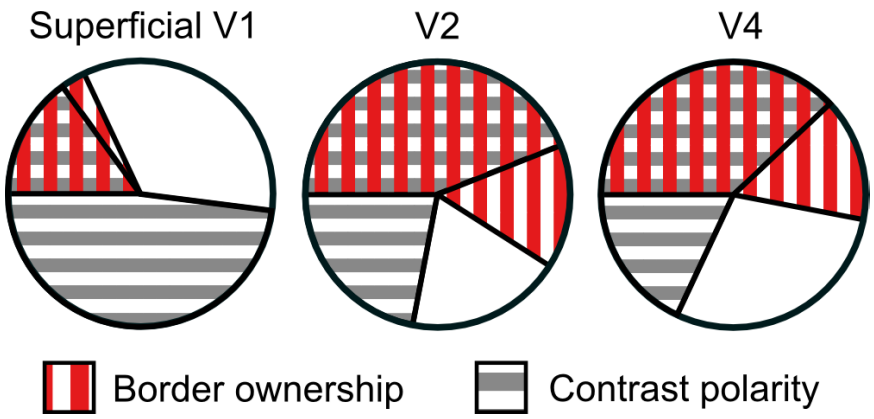


Figure 1.9: Summary of the results from (Zhou, Friedman, and von der Heydt 2000) using the standard (single square) test. Proportions are out of the orientation selective neurons that responded to contrast edges. Neurons that are selective for contrast polarity fire at different rates when the colors within their CRF alternate (for example, a neuron might prefer one side of the border to be darker than the other). Neurons that are selective for border-ownership fire at different rates depending on the location (side of CRF) of the owning object (as in Figure 1.4B). Some neurons respond to both to contrast and border-ownership, indicated with the intersection of the red and gray patterns.

border-ownership signals, they were nearly always in agreement (see Figure 27 of Zhou, Friedman, and von der Heydt (2000)). It was also shown (for two example neurons) that the border-ownership signal was position invariant in the direction orthogonal to the border.

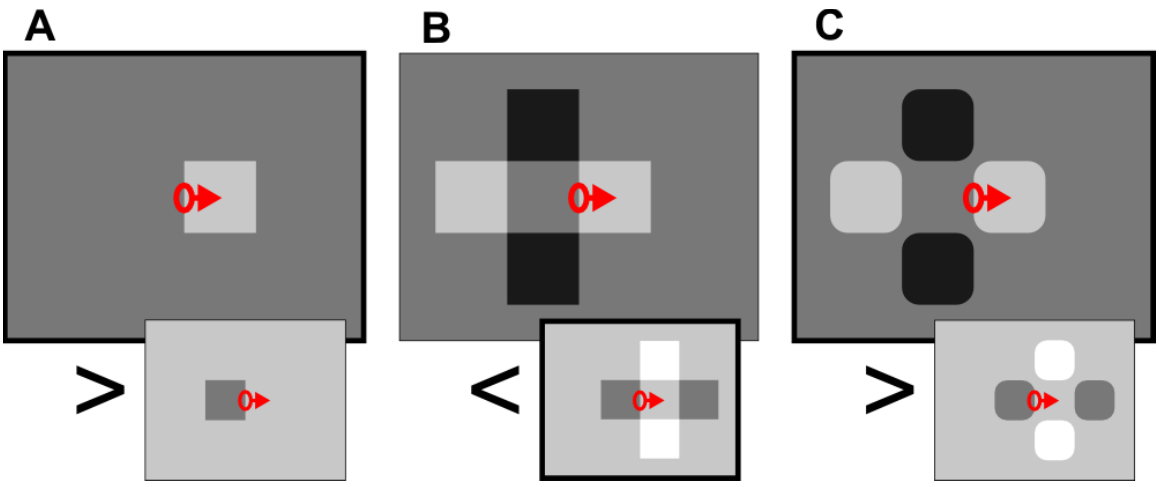


Figure 1.10: Stimuli used in (Qiu and von der Heydt 2007). Border ownership coding was determined by the difference in response between the large displays and the inset displays. The border-ownership signal flips in the middle display (with the inset display eliciting a larger response), in agreement with a transparent overlay interpretation.

Extending the neurophysiological approach to a different perceptual situation, Qiu and von der Heydt (Qiu and von der Heydt 2007) showed that the same neurons also code for border-ownership according to the perception of transparent overlay. When four squares are arranged like in Figure 1.10B, it looks like one semi-transparent bar is overlaying another. With this interpretation, the border in the receptive field (indicated with the red oval) would be owned by the left. If the corners are rounded,

as in Figure 1.10C, the perception is broken and four separate objects are seen and the border is owned by the right. In fact, the average border-ownership signal switched in agreement with the perceptual interpretation.

Zhang and von der Heydt (2010) explored the contribution of individual edges to the border ownership assignment by using contour-defined squares (akin to the Cornsweet illusion figure) and decomposing the contour into fragments. Fragments on the preferred side-of-figure produced facilitation, while fragments on the opposite side produced suppression. The timing of the contributions of the fragments was similar regardless of their proximity to the CRF.

1.2.2 Stereoscopic cues

One of the potential criticisms of the monocular border-ownership tests is that the elicited signals may not be encoding border-ownership. They might, for example, be the result of non-uniform inhibitory or excitatory surrounds that may be unrelated to border-ownership. The object that is closer to the viewer owns the border, because the object must be in front of the background objects in order to visually occlude them. Hence, stereoscopic depth cues are an ideal method of confirming that border-ownership signals are actually coding for border-ownership. This is exactly what was done in (Qiu and von der Heydt 2005), which builds on the findings from using borders defined in random-dot stereograms by von der Heydt, Zhou, and Friedman (2000).

Random-dot stereograms are paired images where the forms of objects are only visible when the images are viewed stereoscopically. Von der Heydt, Zhou, and Friedman (2000) used such stereograms to study the form processing of the supragranular layers of V1 and area V2. Both visual areas contain neurons that respond preferentially to surfaces at a specific depth. However, in area V2, but not V1, neurons were found that responded to borders of figures defined by stereoscopic depth and were tuned to the orientation of the borders, just as they were tuned to contrast-defined edges. Furthermore, most of these neurons also fired at a higher rate when a specific side of the border was closer than the reverse. In other words, these neurons signal the stereoscopic depth order of the surfaces at the edge.

Another study explored the relationship between the border-ownership signal elicited by a solid colored square and the stereoscopic depth order signals (Qiu and von der Heydt 2005). In area V2, 22% of the neurons (37/174) were selective for border-ownership with the contrast-defined figure (without depth) as well as for border-ownership defined by depth in random-dot stereograms (which are devoid of contrast edges). Of this subset, 81% (30/37) had the same preferred side for both stimuli. This correlation shows that the neurons combine different figure-ground cues in a meaningful way. One cue is stereoscopic depth order, the other cue is the global configuration of edges. At contours of occluding objects in the real world, stereoscopic depth is 'near' on the object side relative to the background side. Thus, the observed correlation shows that the visual system treats a figure on a computer display like a

real object occluding a background.

1.3 Persistence and remapping of border-ownership signals

While not directly related to this thesis, there have been some exciting findings about the role of border-ownership in the persistence of visual perception. When looking at visual scenes, humans and many other animals do not maintain a fixed eye position, but continuously make saccades, several times per second. Even though a large part of our visual system is retinotopically organized, we maintain a stable visual perception. O'Herron and von der Heydt (2009, 2011) discovered that the border-ownership signals in V2 neurons often persist for over a second when the figure-ground assignment becomes ambiguous (see Figure 1.11). Even more interesting, this border-ownership persistence can be remapped during saccades and moves with the ambiguous displays if they jump to a new location (O'Herron and von der Heydt 2013). These findings show that border-ownership selectivity reflects a mechanism that helps to maintain a stable visual percept.

O'Herron and von der Heydt (2009) aligned an edge of a square to the CRF at the preferred orientation, as shown in Figure 1.11. A stereoscopic display was used to make the circle appear as a window, with the outside region appearing a few cm in front of the stimuli within the circular window. The square was presented for 500 ms,

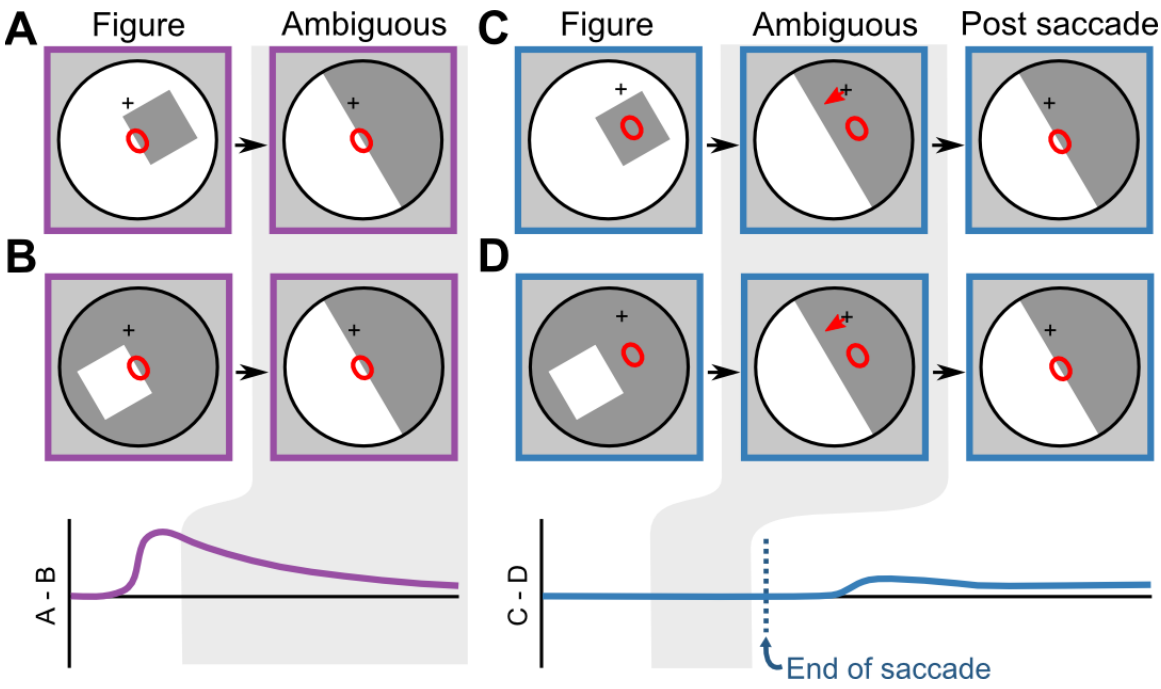


Figure 1.11: Example displays and approximate time course of the population border-ownership signal of experiments by O'Herron and von der Heydt (2009, 2013). In the figure phase of (A,B), a square is shown with one border aligned to a CRF, such that the square is on the preferred (A) or non-preferred (B) side-of-figure. After 500 ms, the display changes to an ambiguous edge. The border ownership signal (A-B) persists, as seen on the peristimulus histograms. In (C,D), a square is shown in the figure phase, but the edges are outside the CRF. After 500 ms, the square changes to a single edge with ambiguous ownership and subsequently the fixation point is moved such that the CRF moves onto the edge. Surprisingly, the border-ownership signal (C-D) appears even in this case where the neuron being recorded is never presented an unambiguous border of the square.

and then switched to an ambiguous display (Figure 1.11A) where the border could be owned by either side. The authors analyzed the border-ownership modulation in the persistence phase (from 200 ms to 1000 ms after ambiguous display onset). Looking at the spike counts during this interval persistence varied a lot between cells, from no persistence to nearly complete persistence. However, the time course of the population signal showed a slow steady decay with a time constant of 400ms.

1.4 Neural models and constraints

As shown in the studies reviewed above, the primate brain has a remarkable ability to calculate border-ownership quickly, even when doing so requires contextual integration over large areas of the visual field. It is a challenge to model how border-ownership coding can be calculated so quickly, considering that the context information is spread out widely in cortex, and neural conduction velocity is limited.

Based on the possible neural mechanisms of propagating context signals across the retinotopic cortical representation one can distinguish three general classes of models: feed-forward, horizontal, and feedback.

1.4.1 Feedforward models

Many neurons have regions outside of their CRF that modulate their response. These modulatory surrounds can be either suppressive or facilitative. Walker, Ohzawa, and

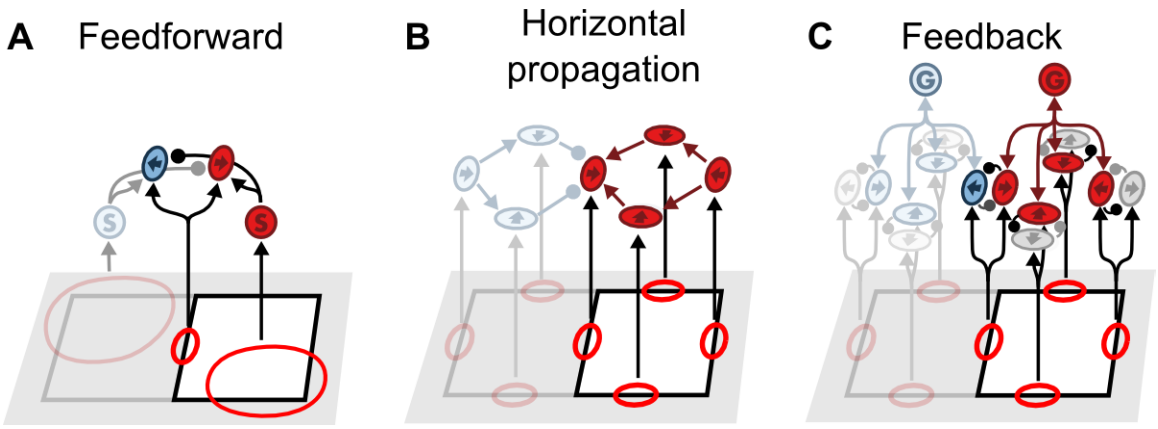


Figure 1.12: Simplified diagrams of an example border-ownership coding model for each of the three general classes. The stimulus display (light gray region, bottom) contains a single white square on the right, with another potential square location outlined on the left. The receptive fields of V1 simple cells are depicted on top of the display with red oval outlines. The vividness of the colors indicate the level of spiking activity. Pointed arrow heads between cells indicate an excitatory connections, while round arrow heads indicate an inhibitory synapse. Colors indicate the preferred figure location on the display (blue: left square, red: right square, gray: other position). (A) A modulatory surround feed-forward model of border-ownership. The "S" cells provide surrounds for the border-ownership cells and either provide facilitation or suppression depending on the side preference. (B) A lateral propagation model of border-ownership. (C) A grouping cell feedback model of border-ownership. The "G" cells are grouping cells that first get excited by the borders that match their preferred figure position and size, and then modulate the border signals, creating border-ownership selectivity.

Freeman (1999) found that the surround regions, measured with grating stimuli, are generally suppressive, and often asymmetric about the CRF. Motivated by these findings, Sakai and Nishimura (2006) showed that a model with asymmetric surround regions (a facilatory region on one side and a suppressive region on the other, Figure 1.12A), stochastically choosen for each neuron, can account in a statistical sense for the data of Zhou, Friedman, and von der Heydt (2000).

Supér, Romeo, and Keil (2010) proposed a feed-forward model that uses two stages of concentric center-surround mechanisms to calculate, first figure-ground modulation, and subsequently border-ownership assignment.

However, the feed-forward models are physiologically implausible. First, the anatomically defined forward connections are precisely what defines the CRF, whereas the non-classical surround is mediated by horizontal connections and feedback from higher areas (Angelucci, Levitt, and Lund 2002).

Second, the cited studies on surround modulation cannot explain the large range of the context influence in border ownership modulation (10 times the extent of the CRF and more). And third, neither of the two model studies addresses the problem of limited conduction velocity and the short latency of border-ownership signals. The virtue of these models is their simplicity, but it is unclear if they can explain critical findings such as the strong border-ownership signals for displays of transparent overlay (Qiu and von der Heydt 2007).

1.4.2 Lateral propagation models

Zhaoping (2005) proposed a model in which border-ownership is calculated within V2, relying on lateral connections (Figure 1.12B). Local borders are represented by two sets of cells, one for each direction of border assignment. The activity of these cells spreads through lateral connections, providing either enhancement or suppression, depending on the shape of the activating contour. By propagating activity along the representation of the contour, the network assigns border ownership to the predominantly concave side of the contour.

It does this correctly even for stretches of contour where the figure is on the convex side as in the case of a C-shaped figure. Sugihara, Qiu, and von der Heydt (2011) argued that such a model cannot explain the short latency of border-ownership signals because of the large distances the signals would have to travel along the contour representation in cortex and the low conduction velocity of horizontal fibers. They measured the latencies of the border-ownership signals for different sizes of squares and calculated the cortical distance to the nearest point of context information. They found that the recorded latencies did not increase as much as predicted by the model.

1.4.3 Feedback models

Feedback models of border-ownership coding (Craft et al. 2007; Jehee, V. A. F. Lamme, and Roelfsema 2007) rely on higher level areas that have larger receptive

fields and modulate the activity in the lower level areas via back projections. Craft et al. proposed a "grouping cell model" in which signals from edge selective cells in V2 are integrated by grouping cells (G) at a higher level. The G cells project back to the same cells they receive input from, facilitating their responses. The G cells have annular integration fields, which makes them most sensitive to compact shapes. For example, when a square figure excites the red set of V2 cells, the corresponding G cell is strongly activated and, by feedback, enhances the responses in the red set, whereas the G cell on the other side receives input only from one edge and is therefore only weakly activated. The feedback makes the V2 cells border-ownership selective.

This model can explain the large context integration and the short latency of the border-ownership signals, because the grouping cells can be in another cortical area so that the feedback signals would travel through white matter fibers which conduct about ten times faster than cortical horizontal fibers (Girard, Hupe, and Bullier 2001). Also, the length of the connections does not increase in proportion to the size of the figure representation in V2 cortex, as would the required length of horizontal fiber connections. This explains the relative invariance of the latency with variation of the size of the squares.

Note that different stimuli may use different types of processing in order to calculate border-ownership. For example, displays in which objects are defined by the configuration of contours may require feedback projections to provide the context information to neurons in V1 and V2, while border-ownership in random-dot stereo-

scopic displays might be calculated in a feed-forward manner.

Another important argument for the grouping cell model is that it can easily be extended to explain selective attention, which is known to spread within objects (Egly, Driver, and Rafal 1994). Because a single grouping cell can facilitate all the feature neurons connected to it, a top-down attention signal only needs to excite a small cluster of grouping cells to enhance the entire contour of an object (Mihalas et al. 2011). A simple consequence of the connection scheme of Figure 1.12C is the asymmetry of the attention effect observed by Qiu, Sugihara, and von der Heydt (2007): A given border-ownership cell, for example the red cell in the center of Figure 1.12C, is facilitated when a corresponding grouping cell is activated, that is, only when a figure on its preferred side of border-ownership is attended. Attention to a figure on the other side does not facilitate the cell, because the grouping cells on the other side project back to the opposing (blue) border-ownership cell.

O'Herron and von der Heydt (2013) showed how this model could be extended to explain the remapping of border-ownership signals.

1.5 Summary of thesis

Chapter 2 gives the general methods used for the thesis and the experimental designs.

Chapter 3 describes the main experiment where natural scenes were used to elicit border-ownership coding in visual area V2. It shows that natural scenes do elicit border-ownership coding, although it is weaker than the border-ownership signal that

is elicited by the standard square stimulus.

Chapter 4 describes additional experiments, in which more complex geometric stimuli and "hybrids" between geometric and natural stimuli (such as silhouettes) were used to study how the different dimensions of scenes affected the border-ownership signal strength visual area V2. It shows that complex geometric stimuli, such as overlapping squares and c-shape figures, also elicit weaker border-ownership signals than the edge of an isolated square, and that "simplified" natural images tend to elicit stronger signals than the original natural images.

Chapter 5 concludes the thesis with the summary of findings and a discussion of how the findings fit into a broader context.

Appendix A shows the results from limited set of neurons that were recorded from visual area V1 (almost all from Monkey M27) and Appendix B shows the results from visual area V2 divided by the monkeys.

Chapter 2

Methods

I recorded from individual neurons in visual areas V1 and V2 of five hemispheres of three alert, behaving male rhesus macaques (*Macaca mulatta*) using extracellular microelectrodes. The distribution of the orientation cells (that were presented with both the standard square and natural image tests) are shown in Table 2.1. This thesis focused on neurons from visual area V2 and the main results are from this area. However, there were some neurons recorded from visual area V1 and these results are given in Appendix A.

Our general methods have been described previously (Qiu and von der Heydt 2005; Zhou, Friedman, and von der Heydt 2000) and are described in more detail in the following sections. All procedures conformed to National Institutes of Health and USDA guidelines as verified by the Animal Care and Use Committee of Johns Hopkins University.

Monkey	Hemisphere	V1	V2
27	Right	33	33
27	Left	3	32
28	Right	1	16
28	Left	0	34
29	Right	1	37
	Total	38	143

Table 2.1: The number of V1 and V2 neurons presented with both the standard square and natural scene side-of-object selectivity tests, by monkey and hemisphere.

2.1 Preparation

The monkeys were first gradually introduced to the lab and the recording room. They were then trained on a joystick task that required them to focus on the monitor for a water or juice reward. Once they were familiar with this task, three small titanium head posts were attached to the skull using short titanium cortical screws and bone cement. The monkeys were put under general anesthesia with subcutaneous injections of atropine (0.1 mg) and then ketamine HCl (20 mg/kg body weight) followed by intravenous infusion of sodium pentobarbital (12.5 mg/kg given initially as bolus and then as needed, at least another 12.5 mg/kg). Buprenorphine HCl was given

post-operatively (0.03 mg/kg) after the monkey began to wake up from anesthesia and again after 12 hours. The monkeys were given at least three weeks to recover following all surgeries.

Stainless steel recording chambers (20 mm × 20 mm, one over each hemisphere) were later implanted over the lunate sulcus to maximize the V2 area accessible from chamber, using the same surgical procedure as above. The chamber was closed with a stainless steel cover with a rubber seal. The bone was coated with Copalite® varnish. On the first day of recording and later as needed, a small trephination (5 mm) was created under ketamine anesthesia, leaving the dura matter intact. All surgeries and procedures involving access to the recording chamber were done under sterile conditions.

2.2 Recording setup

Isolated neuronal activity was recorded extracellularly with high impedance ($\geq 3M\Omega$) glass-coated platinum-iridium microelectrodes ($10\mu m$ exposed tip) that were made in our lab by Ofelia Garalde. The tissue above the dura was thinned before the insertion of the electrode as needed to insure that the microelectrodes could penetrate the dura without bending the tip. Bone wax was placed on the dura to stabilize the dura and the recordings. Dexamethasone was applied to the surface of the dura between consecutive days of recordings to discourage tissue growth.

The signal was amplified 10× by a headstage. The spike signal was bandpass

filtered at 0.25 - 60 kHz (Krohn-Hite) and further amplified and digitized at 50 kHz. Cell spikes were selected by either using a threshold or by using a spike template at the time of the recording using the Alpha Omega spike detection system (ASD 2.80). Only well isolated neurons were recorded.

Many of the neurons, especially in V2, were easily isolated, either because the spikes were from only a single neuron or there were substantial differences in either the recorded spike shape or spike magnitude between the spikes of the neuron of interest and the other spikes.

2.3 Eye tracking

Eye movements were recorded for one eye by corneal reflection or pupil tracking by an infrared video system (Iscan ETL-200) at 60 Hz with a resolution of $5,200 \times 2,560$. The eyes were imaged via a hot mirror. Optical magnification was used, making the resolution of the corneal position 0.08×0.16 degrees of visual angle. The accuracy, however, was limited by noise and drifts of signal.

2.4 Presentation of stimuli

VLAB, a program written by Fangtu Qiu for our lab, was used to control the experiments and to record the spike data. This program ran on a Windows system and controlled the experiments that were displayed by Open Inventor on a Linux system.

The stimuli were presented to the monkeys with either a 21-inch EIZO FlexScan T965 or ViewSonic G220fb color CRT monitor. Each had the refresh rate of 100 Hz and a resolution of $1,600 \times 1,200$. The monitors were about 1 meter away and subtended at least 16×21 deg of the visual field.

The output of the monitors were linearized. Images of natural scenes were adjusted with a gamma value of 2.2 for display on the linearized display.

2.5 Behavioral paradigm

All data were collected using a fixation paradigm, where the monkeys were required to maintain fixation for about 3 seconds for a juice reward. The monkeys would indicate that they wanted to start the trial by looking at a fixation point on the monitor. They had to keep their eye gaze within 1 visual degree of the center of the fixation point in order to receive the reward.

2.6 Mapping of the classical receptive fields

After isolation of a cell, its classical receptive field (CRF) was manually mapped with bars, drifting gratings, and/or rectangles, depending on the CRF properties of the neuron. Different colors and orientations were used to determine the optimal stimulus. The manual mapping was typically confirmed by a position test that systematically moved a bar or step-edge. Other properties of the CRF, such as color and orientation

preference, were also systematically tested.

2.6.1 Orientation mapping

The preferred orientation of each neuron was determined by bars, gratings, and/or squares, depending on the neurons' stimulus preferences. The orientation was first tested in steps of 30 degrees (from 0 to 330 degrees). Sometimes the orientation was fine-tuned using smaller steps over a more focused range.

Cells that were not orientation selective were discarded. This was determined by using the orientation modulation index when orientation tests were available or by manual mapping when the orientation preference was obvious. The orientation modulation index was calculated by:

$$OMI = \frac{R_\theta - R_{\theta_\perp}}{R_\theta + R_{\theta_\perp}}, \tag{2.1}$$

where R_θ and $R_{\theta_\perp}$ are mean responses to the preferred and orthogonal orientations. The preferred orientation was the orientation that was estimated during the recording, which was used for other tests. If the preferred orientation was not one of the orientations that was presented for the 30°-step orientation tests, the nearest orientation was used. The responses to orientations in the opposite direction were combined (i.e. 180 degrees apart). If, for example, the cell was determined to have a preferred orientation of 40 degrees, then the nearest orientation shown during the orientation tests would have been 30 degrees (averaged with the opposite direction of 210 degrees)

and the orthogonal orientation would be 120 degrees (averaged with 300 degrees).

Only cells that had an orientation modulation index of at least 0.20 were included in the analysis.

2.6.2 Position mapping

The position of the CRF was first mapped manually using a bar, grating, and/or square. This mapping was typically confirmed by systematically presenting a bar or step-edge (an edge that bifurcated the screen into two regions of different colors) at different positions orthogonal to the preferred orientation of the neuron and (less frequently) at positions parallel to the preferred orientation. This step was only skipped when the position was clear from the manual mapping.

2.6.3 Color mapping

The color preference of each neuron were typically tested using a bar and/or step edge. The colors tested were the same as Zhou, Friedman, and von der Heydt (2000). Sometimes, if a neuron responded robustly to the current color, the color would not be systematically tested.

2.7 Side-of-object selectivity tests

When the border of a square appears within the CRFs of orientation selective V2 neurons, about half of the neurons will fire at a higher rate when the border is owned by a square on one side of the CRF versus the other side. The side of the CRF that elicits the higher response is called the cell's preferred side-of-square (or, more generally, preferred side-of-object). These cells are said to be side-of-square selective. Previous work has established that side-of-square modulation codes for border-ownership, by correlating the side-of-square modulation with unambiguous cues for border-ownership. For example, Qiu and von der Heydt (2005) showed that border-ownership defined by stereoscopic depth is significantly correlated with the side-of-square modulation and von der Heydt, Qiu, and He (2003) showed that border-ownership defined by motion cutes also significantly correlated with the side-of-square modulation.

In this thesis, I reserve the term "border-ownership modulation" until it is established that the side-of-object modulations are actually coding for border-ownership for a given stimulus type. Side-of-object modulations are modulations in a neuron's response due to the manipulating the position of the object in the stimulus, from the preferred side of the CRF to the opposite side of the CRF. In section 3.3, I show that side-of-object modulations elicited by natural scenes are coding for border-ownership. However, when the global context is taken away from the natural context, and just

a local patch is shown, many cells will exhibit a side-of-object modulation that is not coding for border-ownership. These cells may, for example, be responding to the texture, curvature, or other features within the CRF (in a way that does not utilize natural image statistics to estimate the border-ownership).

The stationary stimuli for all side-of-object selectivity tests were presented for 300 milliseconds, with 100 milliseconds before and after each presentation, making the total inter-stimulus period 200 milliseconds.

2.7.1 Classical side-of-object tests using geometric shapes

Only cells that were tested with the standard square test were included in our analysis. Many of the cells were also presented c-shape figures and overlapping squares. All three of these geometric shapes were used in Zhou, Friedman, and von der Heydt (2000) and are shown in Figure 1.8, page 13.

The standard square test presents a square with one of the borders aligned to the CRF of the cell, with the square rotated such that the orientation of the border matches the cell's preferred orientation. The square and the background were defined by two colors: the color gray and the preferred color of the cell. The assignment of each of these colors to the square and the background were counterbalanced, in order to separate out the effect of local contrast from the border-ownership coding. The

squares were presented in two sizes: 3 or 8°. The color of the screen in inter-stimulus intervals was set to the average of the two colors used for the square and background.

The c-shape figure and overlapping square stimuli were shown to 101 and 122 cells respectively (each of these were also presented the standard test). The c-shape figures and overlapping squares were also presented in two sizes. For the c-shape figures, the mouth or convexity was either 3 or 8° wide, such that the corresponding squares would fit into the c-shape. For the overlapping squares, the occluding square was set to be either 3 or 8°.

The same colors were used for the c-shape as the standard square border-ownership test. For the overlapping square condition, two colors were be assigned randomly to the two squares: the color gray and the preferred color of the cell. The background was then set to the average of these two colors, which is the same as the color of the screen during the inter-stimulus interval for all of the geometric border-ownership tests.

For all three of the geometric stimuli, the following three parameters were systematically tested (the first two depicted in Figure 1.8):

1. **Side-of-object**: the side of the occluding object could be on either side of the receptive field when the edge of the figure was aligned to the preferred orientation of the cell.

2. **Contrast polarity**: the local contrast polarity is the arrangement of the two colors in the CRF. For a given contrast polarity, the stimuli within the CRF

were exactly the same.

3. **Size**: the size of the occluding square (or gap, in the case of the c-shape figure) was set to be either 3 or 8 degrees.

All of the stimuli were presented in pseudo-random order. Initially the batches of the different geometric shapes were tested separately, but were later combined as a single batch containing all three geometric shapes.

Cells had to contain at least one stimulus condition within the standard square stimulus that elicited a firing rate of 5 spikes per second. Cells without a noticeable side-of-object preference with the geometric stimuli were often bypassed, due to the substantial time requirements of the natural scene side-of-object selectivity tests.

2.7.2 Side-of-object selectivity tests using natural scenes

I used training images from the Berkeley Segmentation Dataset 300 (Martin et al. 2001) and one additional image (Neo n.d.). Before starting the recording, I and another lab member, Hee-kyoung Ko, labeled many positions (scene points) on the occlusion-based borders within the images with the orientation of the border and the side of the object that created the border. Custom software written in C++ and OpenCV was used to help reliably define the position of the scene points, orientation of the border, and side of the occluding object.

I used two different criteria for selecting scene points to present to a given cell.

Initially, I only selected scene points where the orientation of the points' corresponding border in the image was within 5 degrees from the optimal orientation of the recorded neuron. In V1, 36 cells (95% of 38 cells) and, in V2, 53 cells (37% of 143 cells) had the natural scenes selected using criteria. Later on, I pre-selected a subset of scene points (from Figure 2.1) and presented all of these scene points to every cell (given that the cell was held long enough), so that the same scene points could be shown across multiple cells and so that I would have time to create high quality "hybrid" versions of each scene point.

Scene points were excluded from the limited set (Figure 2.1) when their object border deviated more than 20 degrees of the tangent line at the scene point within r_{crf^+} pixels of the scene point, where r_{crf^+} is equal to or slightly larger than the radius of the CRF. Additionally, thin objects or parts of objects, such as branches, where you could see both sides of the object within the radius $2r_{\text{crf}^+}$ were excluded. Scene points that were already presented with the most cells took preference in being included in the limited set.

With geometric stimuli, it is easy to control for every factor, such as the side-of-object and the local contrast. This task is much more difficult for natural scenes. This is particularly difficult because the influences with the CRF drive the response and the main interest of in this thesis is the effect of the global context which provides a weaker modulatory influence. Each scene point included the following manipulations (independent variables), depicted in Figure 2.2:

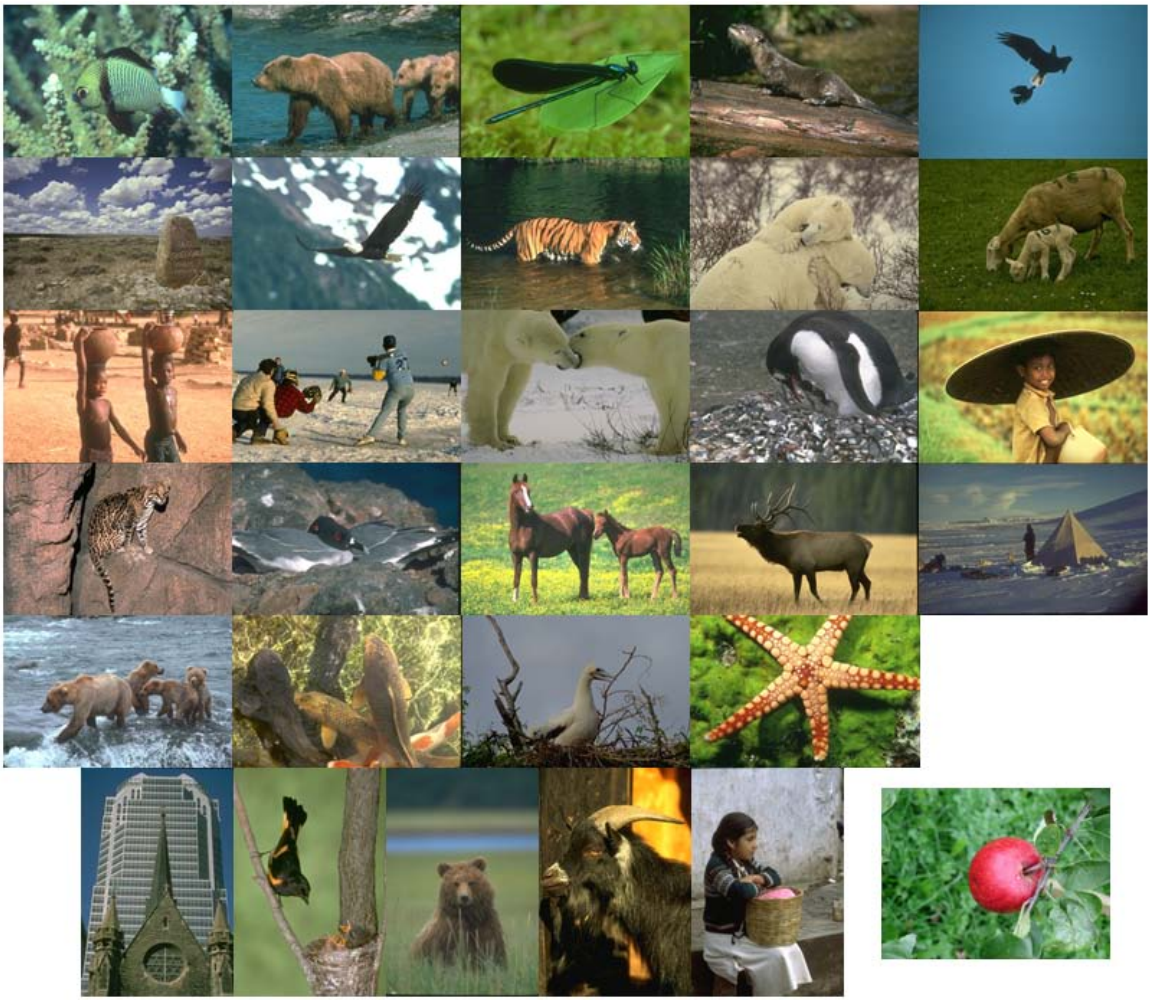


Figure 2.1: Images with scene points used for natural and natural-geometric hybrid experiments. All of the natural scenes shown are from the Berkeley Segmentation Dataset 300 (BSDS300) (Martin et al. 2001), except for the image of the apple on the lower right. Early experiments included many more images from the BSDS300.

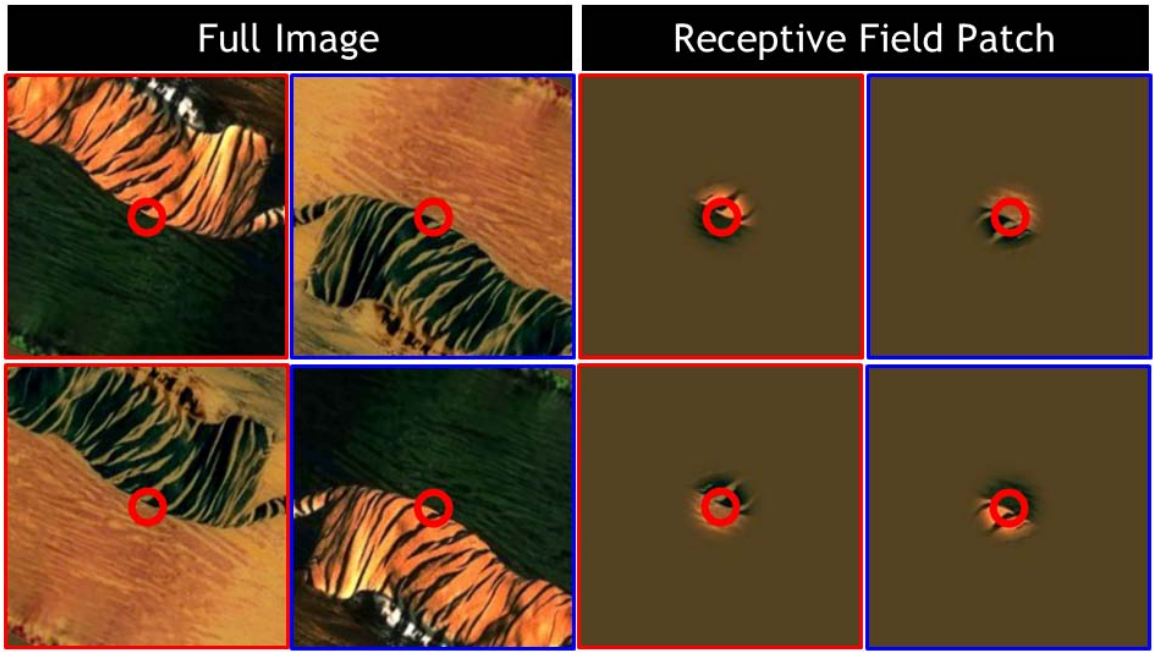


Figure 2.2: All eight of the permutations of the three binary parameters are shown for an example scene point, rotated to match the border in the CRF (red circle) to the preferred orientation of an example V2 cell (28ik1a). The local contrast polarity is the same across each row. The images in the red frames (first and third columns) have the object on to the upper-right of the scene point (which was aligned to the center of the CRFs) while the images in the blue frames (second and fourth) have the object on the opposite side. The side-of-object depends on rotation and the image content: in this image, the most upright natural-color image has the object to the lower-left, however, with other images, the most upright natural color image would have the object to the upper-right.

- **Side-of-object**: The side of the object for a given scene point was manipulated by rotating the images around the center of the CRF by 180 degrees (after the initial rotation to match the orientation of the border to the preferred orientation of the neuron). The side-of-object is then dependent on both the content of the image and the rotation. This parameter was labeled either as the object being on the preferred side-of-square (as determined by the standard square test unless otherwise noted) or as the object being on the non-preferred side-of-square.

- **Contrast polarity**: Rotation of the images switches the colors on the two sides of the border within the CRF. Most V1 and V2 neurons are sensitive to this local contrast polarity (Figure 1.9). In order to control for this, I performed a transformation of the color space that flipped the colors from both sides of the occlusion border of the scene point. The details are given in Section 2.7.2.1.

- **Context**: In addition to showing the full images, local patches were also shown where the context outside of the CRF was removed. This allowed for the response to the local stimulus to be decomposed from response of the global context. The stimulus within a distance of r_{crf^+} from the center of the patch was exactly the same as the full image, where r_{crf^+} was defined to be equal to or slightly larger than the radius of CRF. After the distance of r_{crf^+} from the center of the patch, the image faded into the background with the profile

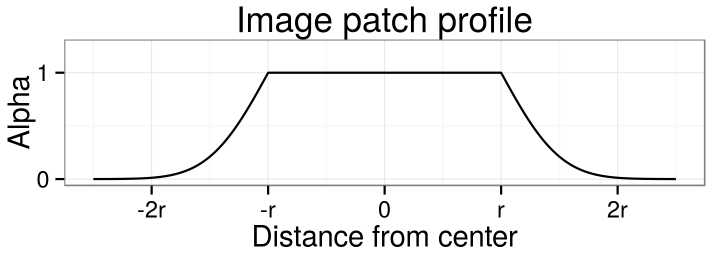


Figure 2.3: The alpha composition profile used for the image patches. This profile was be scaled such that CRF would fit within the $\alpha = 1$ region (the opaque region, $r \equiv r_{\text{crf}^+}$).

shown in Figure 2.3. The value of the radius r_{crf^+} was fixed at 15 pixels and the images were scaled from 1 to 5 arcminutes per pixel. The most frequent scaling used was 2 arcminutes of visual angle per pixel (making the sizes of the images from BSDS300 16×10 visual degrees). Image scaling was performed by Open Inventor.

Each of these factors is binary such that there are a total of 8 variations of each scene point. These stimuli were presented pseudo-randomly in batches of scene points, in order to be able to collect all of the scene points' variants in each batch.

2.7.2.1 Color space manipulation for local contrast

The data from the BSDS300 contained human segmentations of the images (Martin et al. 2001). These segments were used to calculate the average color of the two visual regions on each side of the border at a given scene point.

The average colors were calculated as

$$\vec{c}_{s\in\{fg,bg\}} = \frac{1}{n}\sum_{i\in s}\vec{x}_i, \tag{2.2}$$

for all n pixels $\vec{x}_i = (r_i, g_i, b_i)$ that were in the segment of the border side s (foreground object or background) within the patch radius r_{crf^+} of the scene point (which is aligned to the center of the CRF).

The average of both of these averages was then calculated as

$$\vec{c} = \left(\vec{c}_{fg} + \vec{c}_{bg}\right)/2. \tag{2.3}$$

The color transformation that flips the colors across the border of the scene point was then calculated as

$$f\left(\vec{x}_i\right) = R\left(2\vec{c} - \vec{x}_i\right), \tag{2.4}$$

where R is the rectifying function

$$R\left(\vec{x}_i\right) = \min\left(1, \max\left(0, x_i\right)\right), \tag{2.5}$$

for all pixels $\vec{x}_i$. This flips the color of each pixel around the average color $\vec{c}$. This parameter was labeled by local contrast polarity (i.e. by row in Figure 2.2), analogous with the local contrast polarity of the geometric shapes.

Performing the color flipping procedure twice should make the image approximately the same as the original image. For some scene points, this was not the case. If performing the color flipping procedure the second time did not return the average colors of the foreground object and background to their original values, the scene point was not displayed.

2.7.3 Side-of-object selectivity tests using "hybrid" geometric-natural stimuli

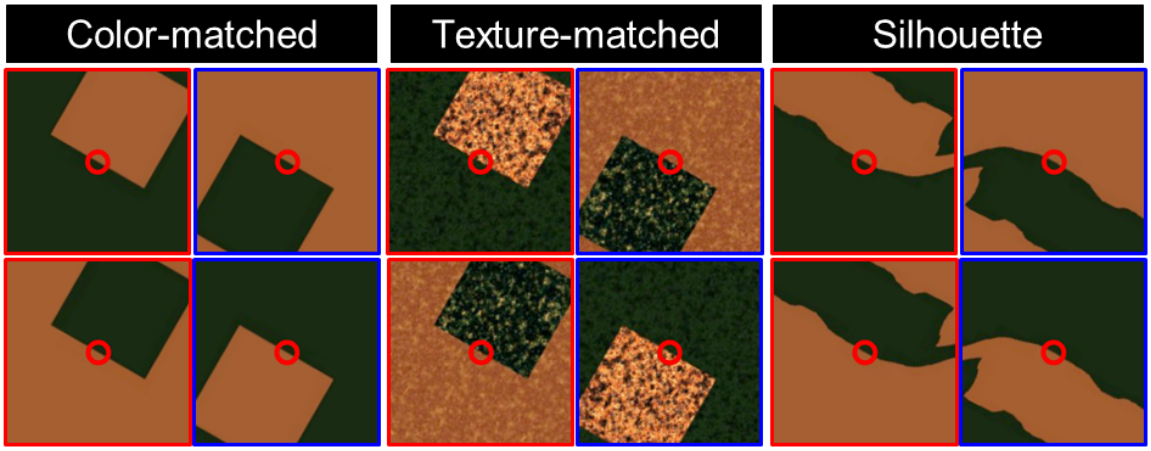


Figure 2.4: Example hybrid versions of tiger scene. The stimuli types are the color matched square, the color and textured matched square, and the color matched silhouettes.

In order to help determine how the border-ownership signals varied along multiple dimensions of natural scenes, I also showed stimuli that were hybrids of the geometric and natural shapes. I used three different types of natural-geometric hybrid versions of the natural scene points (Figure 2.4):

- color matched squares,
- color and texture matched squares, and
- color matched silhouettes of objects.

The first two stimuli types are squares that were defined to be 4 times the size of the estimated CRF diameter ($4 \times 2r$, 4° for the most common scaling). The third type contains the shape (silhouette) of the object whose border falls on the scene point. The average colors of both regions forming the border were calculated from pixels within a distance of r_{crf^+} from the scene point. The average colors of the two visual regions that form the border at the scene point, $\vec{c}_{fg}$ and $\vec{c}_{bg}$, were calculated as described in Equation 2.2 and were used to define the object and background for the hybrid stimuli. The details of the texture synthesis for the color and texture matched squares is given in Section 2.7.3.1.

The silhouettes were manually created around the object of interest using the Gimp software using zoom view and the layer functionality.

2.7.3.1 Texture synthesis

For the color and texture matched squares, textures were derived from both the foreground side and the background side of the border of each scene point. The region used to synthesize the texture for a given side s was restricted to the pixels that were on the human labeled segment of the border side s within the patch radius r_{crf^+} of the scene point, as was used for the average colors $\vec{c}_{fg}$ and $\vec{c}_{bg}$ (Section 2.7.2.1). The limited sizes of the regions made it difficult to replicate the textures well, especially at the lower spatial frequencies. In order to replicate some of the higher level frequencies of the textures within a given image, I first calculated a "difference" pyramid similar

to the Laplacian pyramid for that image, except it contained only three levels and it used linear instead of a Gaussian interpolation. I then use the pixels from this difference pyramid to generate the textures from each side s of the scene point.

In order to make the difference pyramid, I first made a low-pass image pyramid with the original image as the first level, even though only the part around the CRF was used during the synthesis stage (next paragraph). Each successful level was then downsampled, scaling down the image to half the width and height, rounding up the sizes to the nearest integer. Each level i of the difference pyramid was calculated as the difference between level i in the low-pass image pyramid and the level $i + 1$ of the low-pass image pyramid (upsampled to the size of level i, smoothing the image). Three levels were calculated for the difference pyramid.

A new texture pyramid was generated for each side s of the scene point. For each level i, the new pixels were sampled with replacement from the pixels from the corresponding difference pyramid, but only sampling those pixels that were in the segment on the side s and within the scaled distance r_{crf^+} (i.e. $2^{-i}r_{\text{crf}^+}$, where $i = 0$ for the first level) from the scene point. The layers of the new resampled pyramid were scaled to the original size of the image and then summed together, along with the average color $\vec{c}_s$ of that side.

The squares were then formed using the two textures for both the object and background. The texture derived from the occluding object was always applied to the square and the textured derived from the background region to the background. The

normal manipulations were then applied to create the 8 different variations, including the color space manipulation procedure from Section 2.7.2.1. This allowed the colors of the two regions to be flipped while keeping the textures of the corresponding sides, analogous to the natural scenes.

2.8 Data analysis

Data analysis was performed using the R programming language (R Core Team 2013). Plots were made using the ggplot2 package (Wickham 2009).

The spikes between 40 and 300 milliseconds after the stimulus onset were included for analysis. Forty milliseconds was chosen to because V2 neurons begin to respond to the stimulus around this time.

The side-of-object response modulation can be calculated in a way that either is dependent or independent on the firing rate. The post-stimulus histograms (PSTHs) report the side-of-object response difference, defined as:

$$\text{Side-of-object response difference} = R_{\text{pref}} - R_{\text{nonpref}}, \tag{2.6}$$

where R_{pref} is the mean firing rate over the conditions where the object is on the cell's preferred side-of-object (usually as defined by the standard square stimuli) and R_{nonpref} is the mean firing rate over the conditions where the object is on the other, non-preferred side-of-object. This measure is dependent on the firing rate. In order

to show the modulation independent of the firing rate, I also sometimes showed the side-of-object modulation index:

$$\text{Side-of-object modulation index} = \frac{R_{\text{pref}} - R_{\text{nonpref}}}{R_{\text{pref}} + R_{\text{nonpref}}}. \tag{2.7}$$

The definitions of the side-of-object response difference and the side-of-object modulation index are the same for the full images and local patches of the natural scenes as for the geometric tests (Equations 2.6 and 2.7). However, the global context (excluding stimulus local to the CRF) is calculated as the side-of-object response difference (side-of-object modulation index) from the full images minus the side-of-object response difference (side-of-object modulation index) from the local patches. This can be seen as a generalization of the side-of-object modulation index, since the side-of-object modulation from the CRF for the geometric shapes would be zero by definition (except for differences due to trial-to-trial variation).

2.8.1 Analysis of variance

The analysis of variances (ANOVAs) were calculated by the glm and Anova functions (Fox and Weisberg 2011). ANOVAs were performed on square-root transformed spike counts (by the Anscombe transform, which is a square-root transform used to approximate count data with a Poisson distribution as Gaussian). Repeated measures ANOVA was applied on these models to calculate the significance of the effects. The R effects package (Fox 2003) was used to calculate the effect sizes.

A three-way factorial model (side × contrast × size) was used separately for each type of the geometric stimuli. Each of these three factors was binomial.

A four-way factorial model (side × context × scenept × contrast) was used to analyze the natural images. Here, all of the factors are binomial except for the scene point, which can vary between cells (minimum of 10 scene points). The effect of the side-of-object from the global context is interaction between the effects of side-of-object and context. This can be conceptualized as the side-of-object response difference elicited by the full images minus the side-of-object response difference elicited by the local patches (which would be zero if there were no local effects).

A three way factorial model (side × scenept × contrast) was used to calculate the effects of the full images and patches separately.

In order to analyze the effect of images being right-side-up vs upside-down, a three-way model (image orientation × scenept × color mode) was used for the full images (Figure 3.10). The image orientation was a binomial factor for the typical upright or upside-down ($\pm 45^\circ$), with the other rotations discarded from this analysis.

A three-way factorial model (side × contrast × scenept) was used for the hybrid stimuli (Chapter 4). For the cells that were presented with the hybrid stimuli, the same scene points were presented as were presented in the natural scene experiments.

2.8.2 Analysis of neural population

In order to compare the side-of-object effects elicited by multiple types of stimuli across the neural population, I calculated the population fit of the side-of-object effects elicited by two different stimulus types (e.g. standard square and full images of natural scenes) for all cells presented with both tests. The effects were calculated using the method described in Section 2.8.1. The side-of-object effects for all the stimulus types were then normalized by the standard square test: the effects of each cell i were multiplied by w_i, where $w_i = \frac{C}{RSE}$, C is defined such that $\sum_i w_i = 1$, and RSE is the residual standard error (standard deviation of the residuals) from the ANOVA of the standard test.

The preferred side-of-object was defined by one of the two stimulus types being compared, typically the standard square stimuli. The normalized effect for each cell of this stimulus type was plotted along the x-axis, hence all the values on the x-axis are non-negative. The normalized effects for each cell of the side-of-object for the comparison stimulus type (e.g. the natural scenes) were plotted along the y-axis. Cells that have consistent preferences for the side-of-object across the two tests fell above the x-axis, while cells that have inconsistent preferences for the side-of-object fell below the x-axis.

The population fits are calculated using orthogonal regression through the origin. The 95% confidence intervals are calculated by bootstrap (resampling the cells with

replacement), using the R boot package (Canty and Ripley 2014; Davison and Hinkley 1997).

For comparison I also plotted the side-of-object modulation indices from both tests (Equation 2.7). This measurement is independent of the overall firing rate.

2.8.2.1 Time course analysis

For the time course plots (such as Figure 3.5), I took the weighted mean of the PSTHs (1 ms bins) of the neurons, where the weights were the inverse of the standard deviation of the residuals obtained by the ANOVA model of the standard square data. This averaging procedure gives more weight to the cells with less random variation of responses. The time courses were smoothed with a Gaussian kernel with $\sigma = 5$ ms. To make it easier to display, the mean responses and the side-of-object response differences are plotted separately.

2.8.3 Distribution of side-of-object effects across scene points

The methods in the previous section were used to answer whether there is a correlation between the effect of side-of-object between two tests, but does not address how the effect of side-of-object varies between scene points. In this subsection, I first describe how the distribution of effect of side-of-object across scene points for a given cell was

calculated and then how this was done for a population of cells.

2.8.3.1 Within a single cell

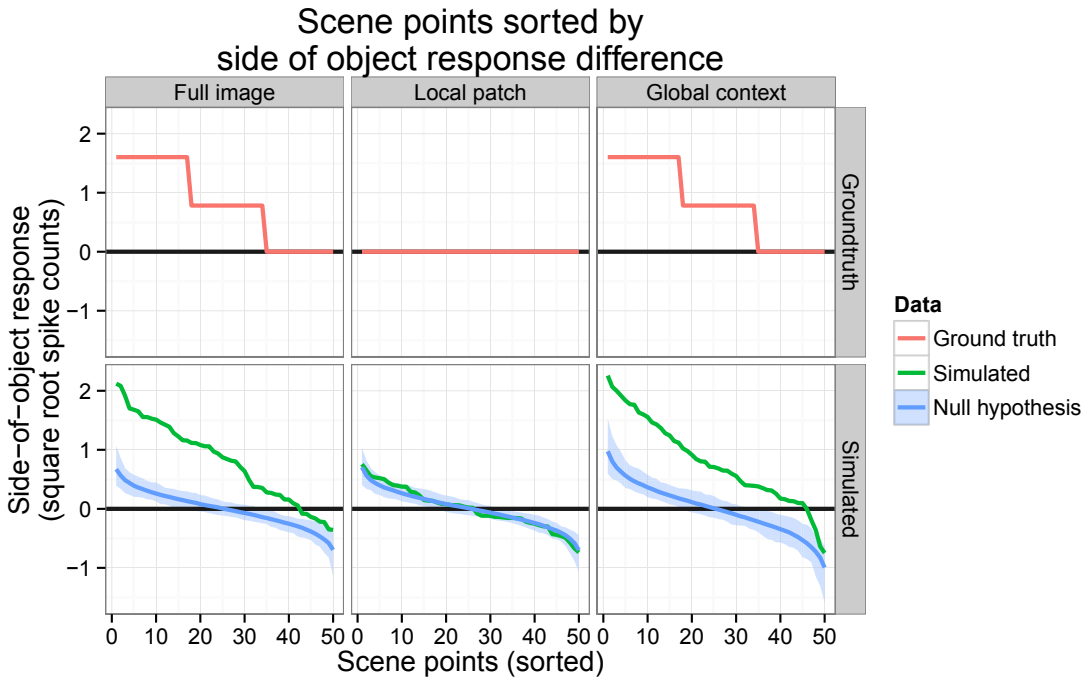


Figure 2.5: Analysis of the side-of-object distribution among scene points for simulated cell. The top graphs show the "true" means of the side-of-object response differences (red). The bottom graphs show the raw scene point data sorted by decreasing effect of effect of side-of-object (green) and the data resampled with replacement irrespective of the condition of the side (blue).

It is difficult to determine how much of the recorded side-of-object modulation for a given scene point is due to trial-to-trial variation and how much is "real" - the

mean side-of-object difference if the cell was presented all of the 8 variations of a scene point an infinite number of times. For Figures 3.6 and 3.7 of Chapter 3, I used the following procedure to plot the distribution of the raw side-of-object effects of the scene points and to approximate the distribution curve for the null hypothesis: that there are no effects of side-of-object (or interactions thereof) and all of the differences in the distributions are the result of trial-to-trial variation.

As an example in Figure 2.5, a side-of-object selective cell is simulated using Poisson distributions. In order to simulate a neuron without any local effects, the spike counts are generated from a single Poisson process with a rate parameter λ, set to 10 spikes per second. For the full image conditions, λ_{pref} and $\lambda_{nonpref}$ were defined for the preferred and non-preferred side conditions such that the side-of-object modulation index would be 0, 0.25, and 0.5 for three separate groups of scene points and such that $\lambda = \frac{1}{2}(\lambda_{pref} + \lambda_{nonpref})$. The "true" side-of-object response differences are depicted in the top row of Figure 2.5.

The side-of-object response difference in the global context of each scene point was calculated as the side-of-object response difference of the full image minus the side-of-object response difference of the local patch of that scene point.

In the first step of the analysis, the side-of-object effects for each scene point was sorted in decreasing order, separately for the full images, patches, and global context. The sorted effects are shown as the green line in the bottom graphs of Figure 2.5. The square root Anscombe transform is used on the spike counts to make the responses

approximately Gaussian, which is needed for generating the null hypothesis.

The null hypothesis, that there is no effect of side-of-object for any scene point, was created by fitting the data with the same linear model used for the ANOVA, setting the main effect of side-of-object and its interactions to zero, and then generating responses by sampling from a normal distribution with the same variance as the distribution of the residuals. This sampling is done many times and the mean is plotted as the blue line and the 95% confidence interval are plotted as the blue shaded region in the bottom graphs. Note that, if there were no trial-to-trial variation within conditions, that this line would be zero.

2.8.3.2 Over cell population

In addition to plotting the distribution of the scene point side-of-object responses for individual cells, I also show the population average of the response distributions in Figure 3.8. Since many cells have been presented a different number of scene points, the response distribution of each cell is linearly interpolated to have the same number of points. For each interpolated point, say the 5th percentile points, the weighted mean side-of-object response of all the cells is calculated. The cells are weighted by the inverse of the standard deviation of the residuals of the natural scene test (normalized so that all the weights add to 1). Hence, cells with a smaller standard deviation of residuals contribute more to the population average.

The null hypothesis distribution is calculated in a similar manner as in Section

2.8.3.1, however, the cells are also randomly resampled with replacement, in addition to the residuals. The linear model for the ANOVA is first fitted to each cell individually and the main effect of side-of-object and its interactions are set to zero. For each sample, the cells are resampled with replacement and the residuals for each presentation are sampled from a normal distribution with the same variance as the distribution of the cell's residuals.

Chapter 3

Border-ownership coding of natural scenes in primate V2

3.1 Introduction

It has previously been shown that when a border of a geometric shape, such as a square, is aligned to the classical receptive field (CRF) of neurons in the early visual cortex of the macaque, that some of these neurons will fire at a higher rate when shape appears on one side of the CRF compared to the opposing side (Williford and von der Heydt 2013; Zhou, Friedman, and von der Heydt 2000). Furthermore, it has been established that these side-of-geometric-object modulations do indeed code for the border-ownership (Qiu and von der Heydt 2005; Williford and von der Heydt 2013). However, it has not yet been shown that natural scenes elicit side-of-object modulations in the early visual cortex or that these side-of-object modulations code for border-ownership in natural scenes.

In this chapter, I show that a subset of neurons in visual area V2 are significantly

modulated by side-of-object in natural scenes: i.e. when the border of an object within a natural scene overlaps the CRFs of neurons in V2, many of these neurons will fire action potentials at a higher rate when the object creating the border is located on one side of the CRF (the preferred side-of-object) versus the opposite side. Furthermore, I demonstrate that the side-of-object modulation elicited by the global context of natural scenes codes for border-ownership by showing that there is a significant consistency between the side-of-object selectivity produced by the standard square stimuli and the side-of-object selectivity produced by the global context of the natural scenes.

When using natural scenes, it is impossible to control for every possible factor, such as contour curvature, texture, and shading. Manipulating side-of-object in a natural scene could modulate a neuron's response without necessarily being the result of mechanisms related to the coding of border-ownership. I therefore make the distinction between the effect of side-of-object, by which I mean any response modulation from manipulating the side-of-object, and border-ownership coding, which implies a modulation that encodes the border-ownership.

In this thesis, I reserve the term "border-ownership signal" for side-of-object modulations that are shown to encode border-ownership. This is established by showing that the side-of-object differences from natural scenes (or, in the next chapter, other stimuli) are correlated with the standard square test, since the standard square test has already been shown to code for border-ownership (discussed next paragraph).

CHAPTER 3. BORDER-OWNERSHIP CODING OF NATURAL SCENES IN PRIMATE V2

This chapter shows that the side-of-object response differences from the global context of the natural scenes correlate with the standard square test, but fails to find a relationship between the local patches and the standard square test (Figure 3.3). Hence, while the local patches of the natural scenes elicit a significant side-of-object effect in many of the cells, these side-of-object effects appear to be encoding local cues (such as texture and curvature) that do not provide information about the location of the occluding object.

An object border is always owned by the object that is closer to the viewer, since closer objects visually occlude objects that appear behind it. Using an edge that is defined by two surfaces at two different stereoscopic depths, some cells in V2 will fire at a higher rate when the surface of one side of a edge appears closer than the opposing surface. Therefore, one way to show that the side-of-object response modulation codes for border-ownership is to show that the neurons have the same preferred side-of-object as it does for the preferred side for the *closer surface*, as defined by stereoscopic cues. Our lab has previously shown that when border-ownership cells in V2 have a preferred side for the closer surface, this preferred side significantly correlates with the preferred side-of-square (Qiu and von der Heydt 2005). This indicates that the side-of-object response modulations elicited by the standard squares codes for border-ownership modulation. In this chapter, I compare the effects of side-of-object elicited by multiple natural scenes with the effects of side-of-square elicited by the standard squares in order to establish border-ownership coding in natural scenes.

3.2 Side-of-object selective cells

Cells were tested for side-of-object selectivity using the standard square test before being presented the natural images. Typically cells without a noticeable side-of-object selectivity where bypassed, hence the population of cells present with both the standard square and natural scene stimuli has a higher proportion of side-of-object selective cells than V2 in general, which is typically around half for the standard square test. A total of 143 V2 cells were tested for an effect of side-of-object using both the standard square and the natural scene tests (with two or more presentations of at least 10 scene points). Of these 143 cells, 89 cells (62%) had a significant ($p < 0.01$) side-of-square effect and 66 cells (46%) had a significant effect of side-of-object with the full images. Figures 3.1 and 3.2 show examples of cells that had significant side-of-object effects with both the standard square and full images. However, the manipulations to the full images do not by themselves completely control for influences within CRF, which drives the main response of the cells. When the global contexts of the natural scenes were removed and only the local patches around the CRFs remained, 36 cells (25%) showed a significant effect of side-of-object, including the example cell in Figure 3.1. When taking into account the local patches (Section 2.8.1), 39 cells (27%) had a significant effect of side-of-object with the global context of the natural scenes.

There were 34 cells that had a significant side-of-object effect with both the stan-

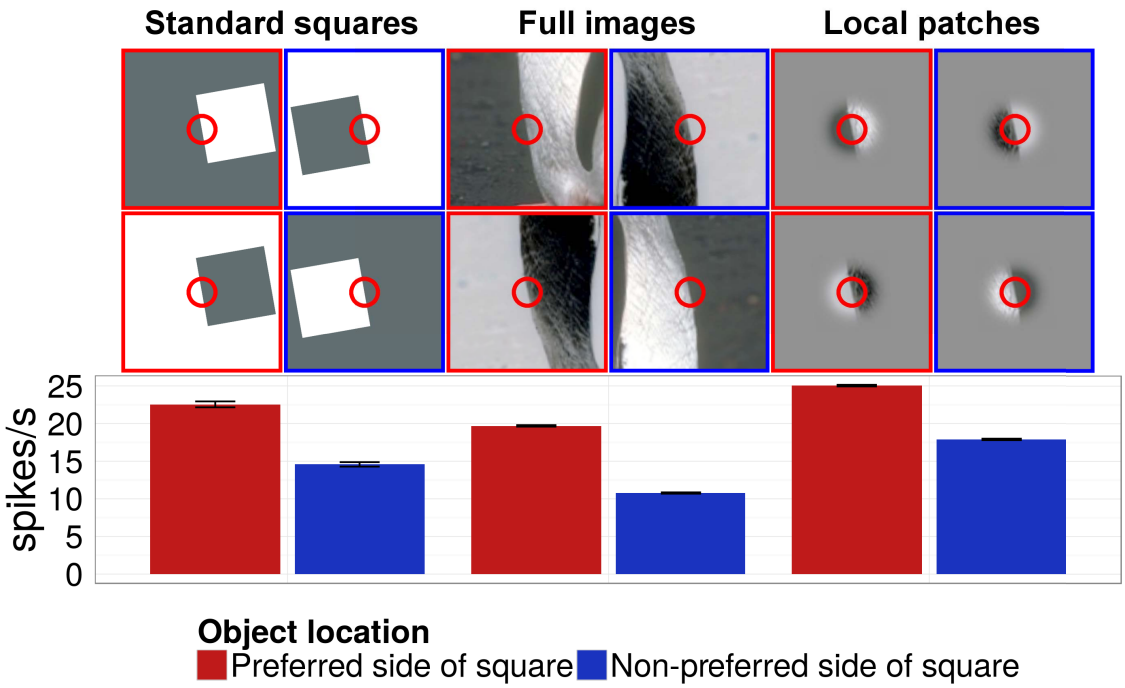


Figure 3.1: Responses from cell 27lj2d to standard square and natural scenes. Only a small window of the entire stimulus screen is presented for each condition (about 7×7 degrees). One of the 177 scene points shown to this cell is depicted for the natural scenes. The CRF is indicated by the red circles (not shown to subjects). The red frames and red bars indicate the stimuli where the object is on the example cell's preferred side-of-square (to the right of the CRF). Blue indicates when the figure is on the opposite, non-preferred side. The side-of-object response difference for each stimulus type is the difference between the red bars and the blue bars. This cell shows a selectivity of side-of-object for the full images and the patches. The preferred side-of-object for both the full images and the patches is consistent with the preferred side-of-square. The whiskers indicate the standard error of the mean.

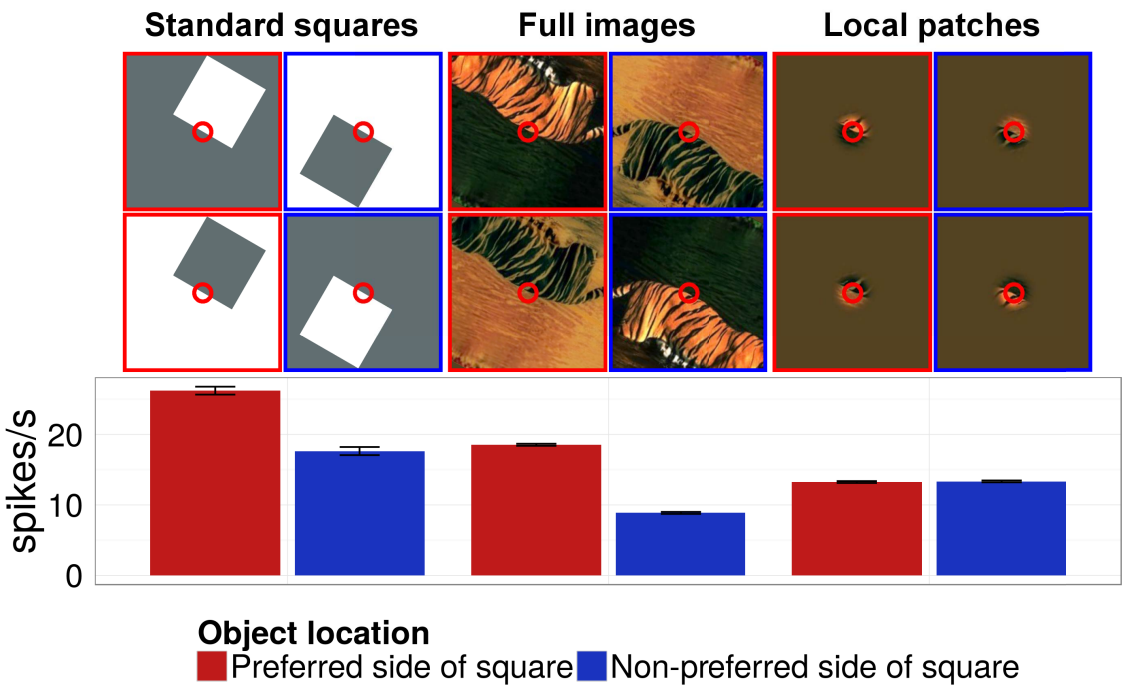


Figure 3.2: Responses from cell 28ik1a to standard square and natural scenes. This cell shows selectivity of side-of-object for the full images, but not the patches. The preferred side-of-object from the full images is consistent with the preferred side-of-square. Only 1 of the 177 scene points presented to the cell is depicted.

dard square test and a significant side-of-object effect with the global context of the natural scenes. Of these, 32 cell had a consistent preference for the side-of-object between the two tests (94%).

3.2.1 Example cells

Two example side-of-object selective cells are shown in the Figures 3.1 and 3.2. The preferred and non-preferred side-of-square (defined by the standard square test) is denoted as red and blue, respectively, for both cells. The cell in Figure 3.1 had a significant side-of-object effect in both the full images and the local patches, which can be seen in the higher firing rate when the objects appear on the preferred side-of-square (stimuli in red frames) than when the objects appear on the non-preferred side-of-square (stimuli in blue frames). However, the side-of-object effect from the local patch may be more of a consequence of uncontrolled local factors than border-ownership, as suggested by the population results presented below (Figure 3.3b).

The cell in Figure 3.2, cell 28ik1a, was more similar to the population average result, where the side-of-object response difference can be seen in the full image displays, but not the local patches. The side-of-object difference of the global context can be thought of as how the side-of-object difference changes from the local patch display to the full image display (side-of-object difference of full image minus that of local patch display). For the standard square test, the local patch response would be zero on average, since the local stimulus is kept identical across the conditions

of side-of-object (Figure 1.8). In the ANOVA, the side-of-object difference of the global context is represented by the interaction between the side-of-object and context variables.

3.3 Border-ownership coding: consistency of side-of-object with standard test

An important question is whether the side-of-object modulations in the natural scenes code for the border-ownership. In other words, are the preferred side-of-object from the natural scenes consistent with other stimuli types or across multiple natural scenes? I first determine whether whether there is consistency between the side-of-object modulation in the natural scenes and the side-of-square modulations in the standard square test. In order to compare the side-of-object effects of these two tests, I first plotted the effects for each test along separate axes in Figure 3.3 for all of the 143 V2 cells. Both effects were normalized by dividing the coefficients by the residual error obtained from the standard square test (see Section 2.8.2). This reduces influence of cells with high unexplained variance while normalizing the effects from both tests by the same amount. I then performed bootstrap on the slopes of the orthogonal least-squares regression lines to calculate their 95% confidence intervals (see Section 2.8.2). The results show that there was a significant agreement between the side-of-object effect elicited by the full images with the standard squares and

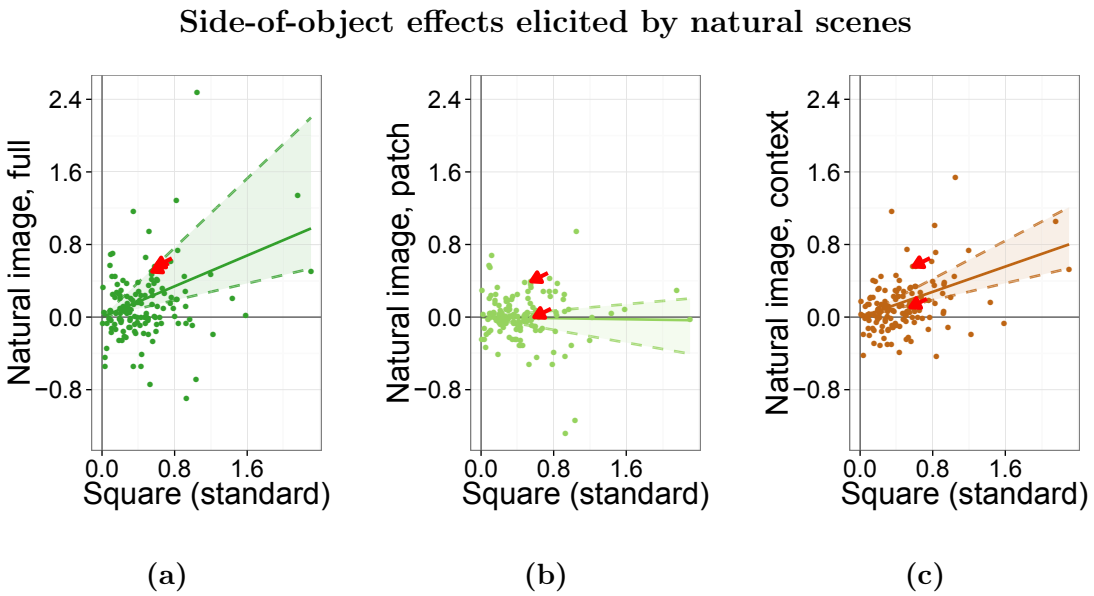


Figure 3.3: Population comparison of side-of-object effects between standard square and natural stimuli. The side-of-object effect elicited by the standard square test is plotted along the x-axis and side-of-object effect elicited by a) full images, b) CRF patches, and c) global context are plotted along the y-axis for all the V2 cells presented with at least 10 scene points (with all eight variations presented at least twice, 143 cells). The solid lines indicate the population fits and the shading and dashed lines indicate the 95% confidence intervals. The example neurons 27lj2d and 28ik1a are indicated with red arrows. The two arrows overlap in (a).

the global context with the standard squares. The local patches failed to show a significant agreement with the standard squares.

Each cell's preferred side-of-square was used as the preferred side in the calculation of the border ownership signal, such that the values on the x-axes are non-negative. The side-of-object effects for the natural images were plotted along the y-axes. Cells that preferred the side-of-object in natural scenes to be on the same side as the standard square fall above the x-axis, while cells that have an inconsistent preference of side-of-object and side-of-square fall below the x-axis. The population fits were calculated using orthogonal regression through the origin (the line obtained by minimizing the squared orthogonal distances to the points is independent of which test is used to define the preferred side-of-object and which test is plotted on which axis). The slope of the population fit for the natural scenes was 0.42 for the full images, -0.01 for the local patches, and 0.35 for the global context. The 95% confidence intervals were 0.23 to 0.95 for the full images, -0.18 to 0.09 for the local patches, and 0.23 to 0.53 for the context images.

The results were nearly identical when the side-of-object modulation indices were used in the above procedure instead of the normalized side-of-object effects (Figure 3.4). The slopes of the population fits using the side-of-object modulation indices were 0.43 (0.23 to 0.7), -0.06 (-0.18 to 0.04), and 0.39 (0.26 to 0.52) for the full images, local patches, and global context.

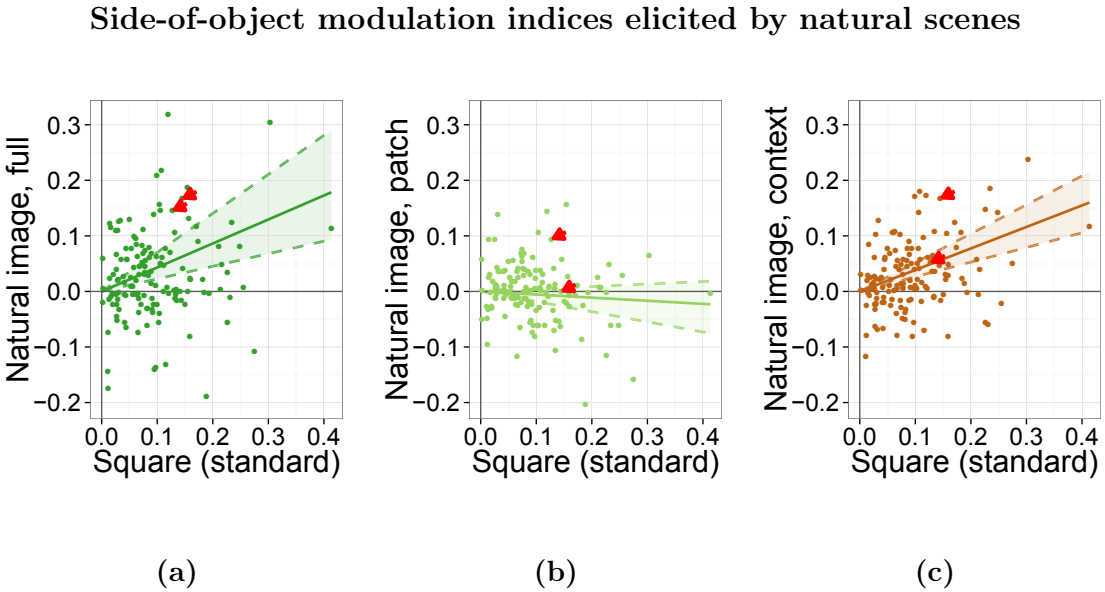


Figure 3.4: Same plot as Figure 3.3, except with the modulation index plotted instead of the normalized effects.

3.4 Time course

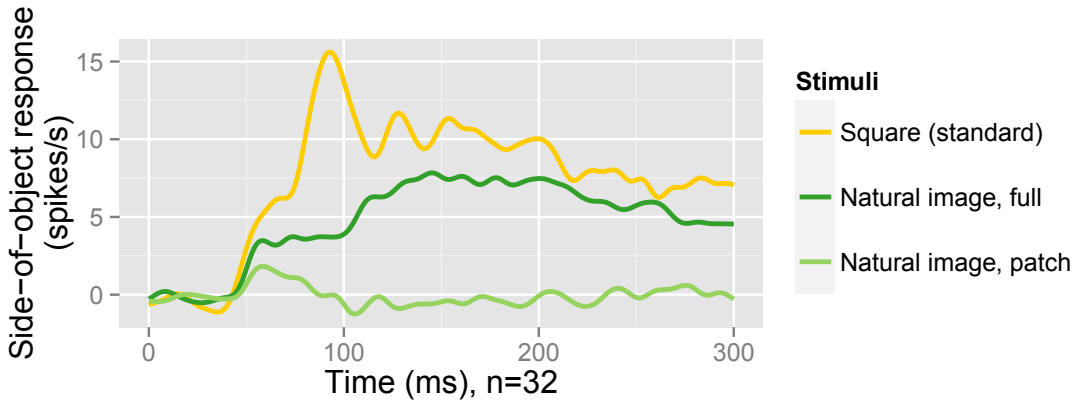
I computed the post-stimulus histograms (PSTHs) for each side-of-object and context condition in Figure 3.5. Only cells that had a significant effect of side-of-object ($p < 0.01$) for both the standard test and the global context of the natural scenes were included. The cells that had an inconsistent side-of-object preference between the standard test and the full and global context were excluded, in order to avoid the bias of using one of the tests to define the side-of-object preference for the other test. There were 32 cells that fulfilled both criteria. Of the cells that had a significant effect of side-of-object for both the standard test and the global context of the natural scenes, only two cells had inconsistent side-of-object preferences.

The mean response and side-of-object response differences are plotted as function of time after the onset of the stimulus. While signals from both the standard squares and the global context of the natural scenes appeared quickly after stimulus onset, the signal for the natural scenes was slower.

3.5 Side-of-object responses across scene points

How does the effect of side-of-object vary across scene points in a given cell? This is shown for the example cells 27lj2d and 28ik1a in Figures 3.6 and 3.7. The scenes were sorted by decreasing effect of side-of-object, separately for the full images, local

(a) Mean response



(b) Side-of-object modulation response

Figure 3.5: Time course of responses elicited by standard squares and natural scenes. The mean responses are shown on top and the side-of-object response differences (Equation 2.6) are shown below. Cells were selected for having a significant effect of side-of-object in the standard test and significant effect of side-of-object in the global context of natural scenes ($p < 0.01$) and for having a consistent side-of-object preferences (32 cells).

patches, and global context. The shaded region is the 95% confidence interval of the null hypothesis: that there is no effect of side-of-object in any of the scenes, or interaction of side-of-object with any of the other conditions. The deviations of the data line from the null hypothesis line show that some of the variance in the border-ownership signal is due to the differences between scenes rather than just random response variations. In example cell 1, most of the variance from the scene differences is due to variation of local cues, that is, features within the CRF (Figure 3.6). In example cell 2, the variance comes almost entirely from the global context (Figure 3.6).

Figure 3.8 shows the population responses averaged over cells with a significant effect of side-of-object ($p < 0.01$) for both the standard square and the global context of the natural scenes and that were presented at least twice with 20 scene points in all 8 conditions (30 cells). The preferred side-of-object was defined by the square test. Therefore, if the effects of side-of-object contained no border-ownership and instead the effects of side-of-object were just from uncontrolled factors that did not correlate with actual border-ownership, the response distribution line and the null hypothesis line would be expected to intersect close to the 50% percentile. There was significant variation in the side-of-object responses elicited by the local patches, however, the lines intersect near the 50% percentile, suggesting that there is no consistent border-ownership signal in the side-of-object response differences elicited by the local patches. The full image data includes the variance induced by features in the CRF. This is

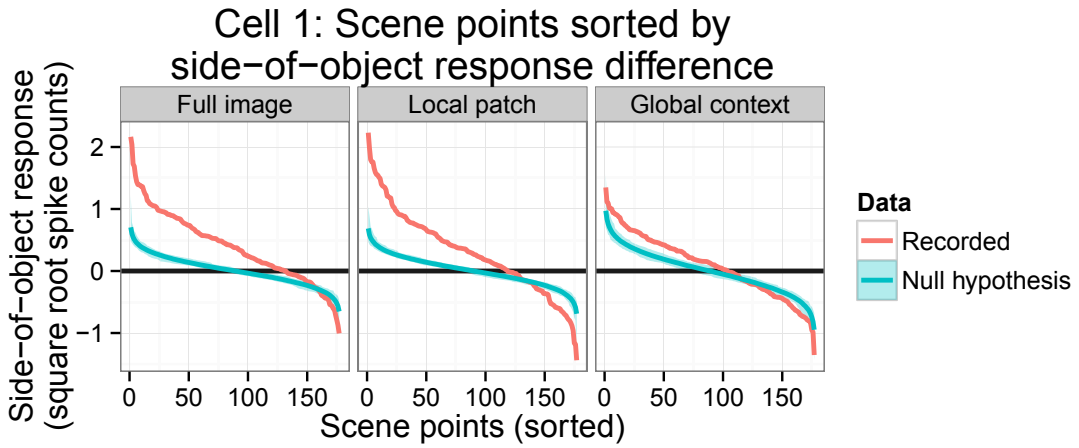


Figure 3.6: Scene points sorted by side-of-object response difference for cell 27lj2d. The null hypothesis line indicates the expected line from sorting random, trial-by-trial variations. Almost all of the scene points in the full images elicited response differences greater than the null hypothesis line, but there were also scene points that appear below the null hypothesis line. The plots show that some of the variance in the border-ownership signal comes from the differences between scenes rather than being unexplained trial-by-trial variance.

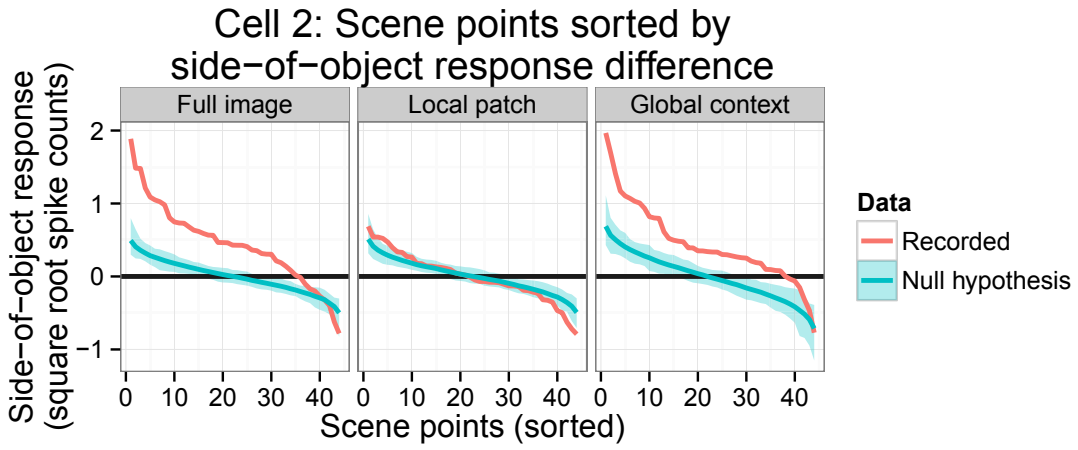


Figure 3.7: Scene points sorted by side-of-object response difference for cell 28ik1a. This cell showed little contribution of scene differences within the CRF, with most of the side-of-object response differences coming from trial-by-trial variances. Almost all of the scene points produced border-ownership signals in the full images and most of this was caused by the global context.

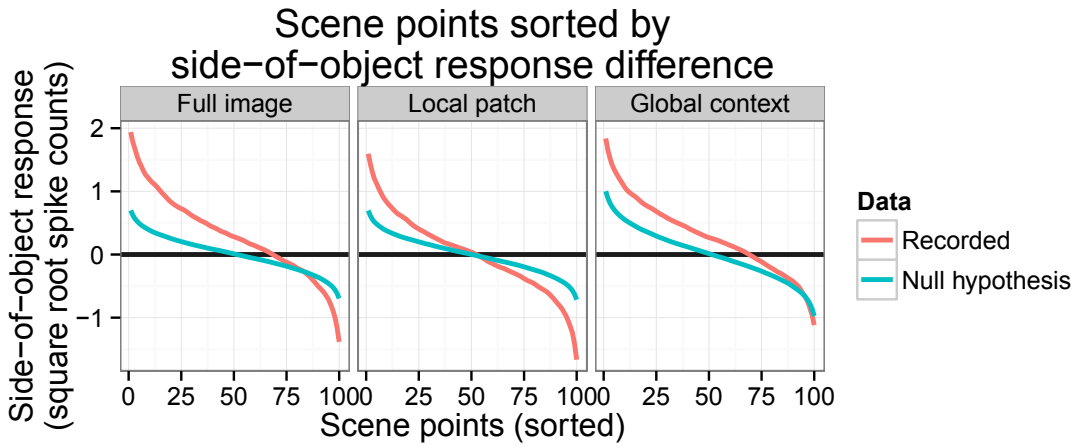


Figure 3.8: Scene points sorted by side-of-object response difference for the population of 30 V2 cells with significant side-of-object selectivity.

presumably why about 10% of the scenes have negative effects (compared to the null hypothesis) in the full image data, whereas hardly any scenes have negative effect from the global context.

3.6 Surround suppression uncorrelated with side-of-object modulation

Willmore, Prenger, and Gallant (2010) used natural images to study the spatiotemporal receptive field of V2 neurons. They found that the distribution of V2 neurons fell into two main clusters: 43% of the neurons fall into cluster A, which have little to no suppressive tuning, and 57% fall into cluster B, which have substantial suppres-

sive tuning. Only 22% of V1 neurons fell into the type B cluster. These ratios are remarkably similar to that of border-ownership selective neurons in both V1 and V2 found in Zhou and von der Heydt (2000), where about 18% of supragranular V1 and 59% of V2 neurons had border-ownership selectivity. Additionally, our lab has found that suppression when the figure is on the non-preferred side tends to be stronger than the enhancement when the figure is on the preferred side (Zhang and von der Heydt 2010).

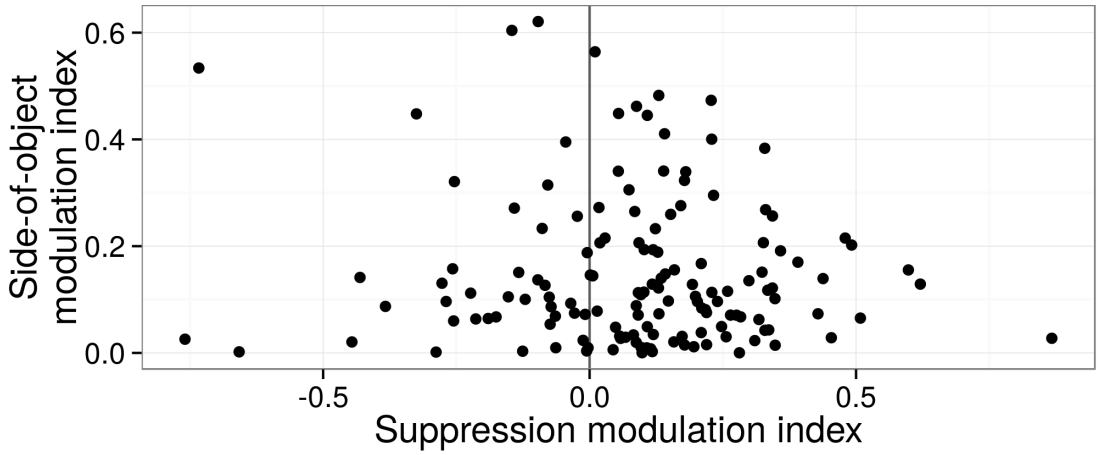


Figure 3.9: Comparison of surround suppression and the side-of-object modulation produced by the global context of natural scenes. There was no significant positive correlation between the surround suppression and the side-of-object modulation (Wilcoxon signed-rank test).

I hypothesized that neurons with more border-ownership coding would be more likely to belong to the class of neurons with suppression (type B). The clustering

analysis used in Willmore, Prenger, and Gallant (2010) requires the presentation of more natural images than what I presented in this study. Instead, I used the suppression modulation index ($(R_{full} - R_{patch}) / (R_{full} + R_{patch})$) and compared this index with the border-ownership modulation (Figure 3.9). There was no significant positive correlation (the Pearson correlation coefficient was -0.077). There was also a lack of positive correlation between the suppression modulation index and the side-of-object modulation index as elicited by the standard square test (not shown). This suggests that the overall surround suppression is not related to with border-ownership coding.

3.7 Effect of image rotation on neural responses

Since I determined border-ownership modulation using "color-flipped" and rotated images, a concern is that the cells might fire at a higher rate when the images are displayed upright compared to when the images are displayed upside-down. However, there was no significant effect across the population for the upright versus upside-down images, as determined by the Wilcoxon signed rank test ($p < 0.05$). The histogram of the image orientation effects of the V2 cells is shown in Figure 3.10. The effects were normalized by the inverse standard deviation of the residuals.

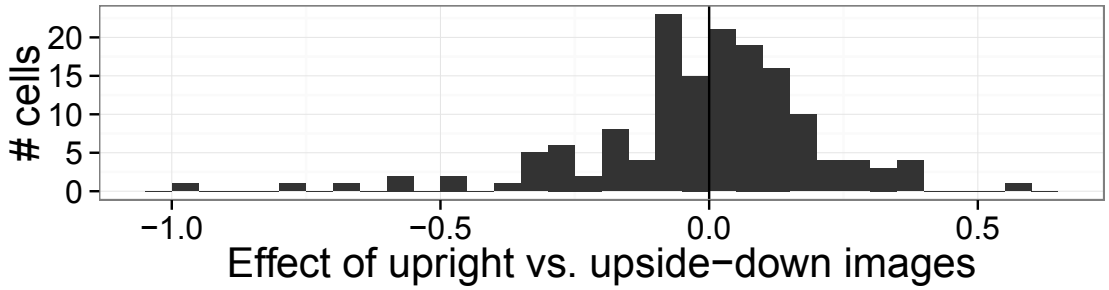


Figure 3.10: Effect of image orientation on neural responses. Histogram of the cells showing the standardized effect of the image orientation. The x-axis shows the standardized effect of displaying the full images upright (±45°) compared to displaying them upside down (±45°). The neurons with a positive value fired at a higher rate for the upright images compared to the upside-down images. There was no significant effect of displaying the images right-side up or upside-down, as determined by the Wilcoxon signed-rank test ($p = 0.33$).

3.8 Conclusion

This chapter showed that natural scenes do elicit a significant side-of-object response in a subset of cells in visual area V2 of the macaque. In the population of neurons that were presented with a sufficient number of the natural scenes, 66 of 143 cells (46%) had a significant effect of side-of-object in the full image condition.

I also show that this side-of-object effect reflects the border-ownership signal, as indicated by the consistency of the preferred side-of-object between the standard square stimuli and the global context of the natural scenes (Figure 3.3) and between different natural scenes (Figures 3.6-3.8). There was no significant consistency between the standard square stimuli and the local patches that covered the CRFs.

While the local patches sometimes elicited an effect of side-of-object, there was not a significant correlation with the effect of side-of-object elicited by the square across the population. If the CRFs in these V2 cells contribute at all to the border-ownership signal, that contribution must be weak compared to the border-ownership signal elicited from the global context of the natural scenes.

The border-ownership signal from the global context of the natural scenes appeared later after the onset of the stimulus compared to when a single square stimulus was presented. However, this signal still appeared before 75 ms and plateaued as early as 150 ms after stimulus onset.

Chapter 4

Border-ownership coding of geometric, natural, and "hybrid" stimuli in primate V2

The previous chapter demonstrated that the global context of natural scenes elicits border-ownership coding in the macaque visual area V2. However, the border-ownership signal was weaker for the natural scenes than the standard square. In this chapter, stimuli that are "hybrids" between natural scenes and geometric stimuli were used in order to evaluate how different dimensions of the stimuli affect the border-ownership signal. Two additional geometric stimulus types were also presented.

This chapter includes the results collected from four hemispheres of three monkeys (M27, M28, and M29). These results show that border-ownership signal strength decreases with the complexity of the stimulus. This is first shown using geometric shapes: the standard square elicited stronger border-ownership signals on average than the c-shape figures and overlapping squares (Figure 4.3). Additionally, the natural scenes elicited significantly weaker border-ownership signals on average than the simplified silhouette versions of the same scenes (Figure 4.9).

There was a borderline significant decrease in the border-ownership signal when either the natural color was applied to the squares or both the natural color and texture of a scene was applied to a square (Figure 4.5 and 4.7). However, this may partially be due to a reduction in overall firing rate from the natural color contrasts. When the modulation index was used, the reductions in border-ownership signal strength were no longer significant (Figures 4.6 and 4.7).

Applying the shape of the natural object to the square did not significantly change the border-ownership signal strength (Figure 4.8).

4.1 Example cell

The geometric stimuli used for example cell 28ik1a and its responses are shown in Figure 4.1. Both the side-of-object response difference (the difference between the red and blue bars, equation 2.6) and the side-of-object modulation index (equation 2.7) were weaker for the c-shape figure and overlapping squares than the standard square. Hybrid versions of many scene points (at least 10) were presented to each cell. The hybrid versions for one of the scene points presented to the same cell are shown in Figure 4.2. All of the stimulus types elicited a consistent side-of-object preference: the upper-right of the classical receptive field (CRF).

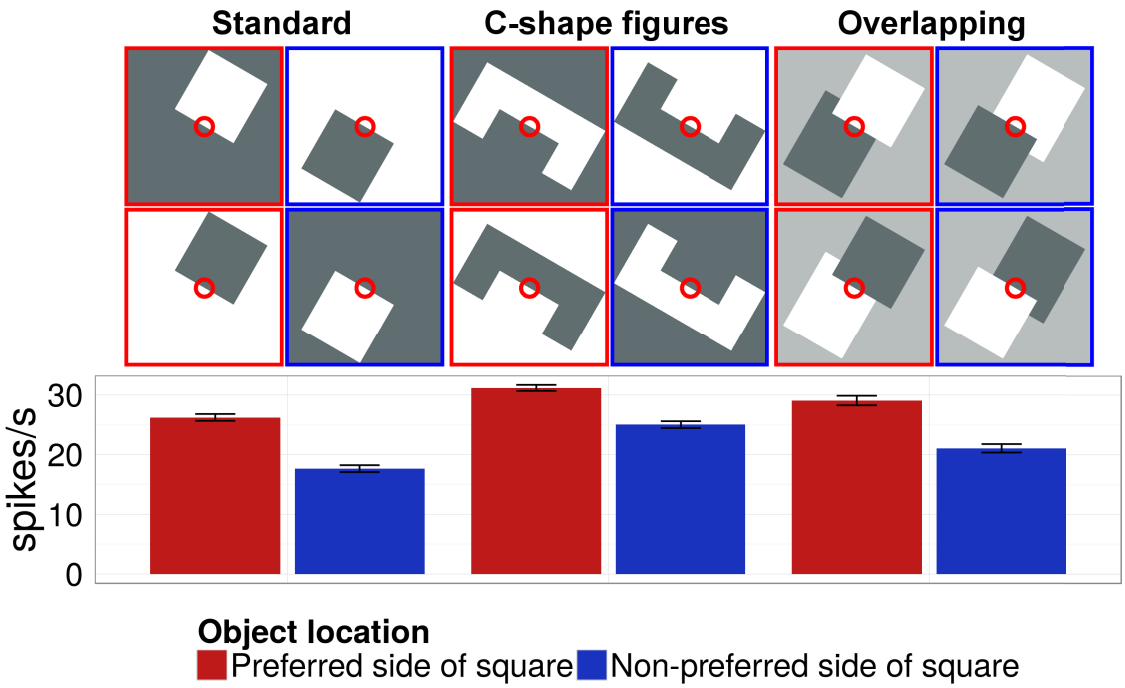


Figure 4.1: Responses from V2 cell 28ik1a to geometric stimuli. The four variants of each geometric stimulus type are shown: square, c-shape figure, and overlapping squares. This cell shows the same preferred side-of-object for all three stimulus types.

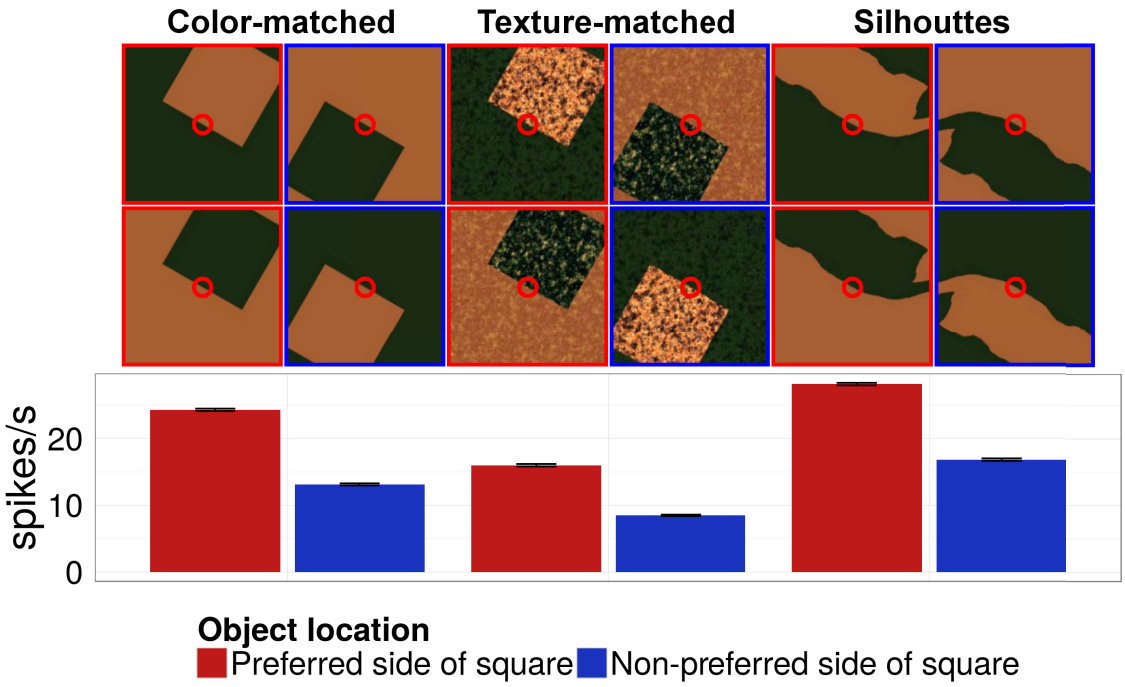


Figure 4.2: Responses to natural-geometric hybrid stimuli from V2 cell 28ik1a, the same cell as in Figure 4.1. The four variants of the tiger scene point are shown for the three types of hybrid stimuli: color-matched square, color and texture matched square, and silhouette. This is the same scene point that was presented in Figure 3.2. A different set of hybrid stimuli is derived separately for each scene point. This cell shows the same preferred side-of-object for all three hybrid stimuli, in addition to the geometric stimuli. Only a small window of the entire stimulus screen is presented (about 7×7 degrees).

4.2 Stimulus comparisons across neural population

4.2.1 Responses to geometric shapes

In order to understand the differences between the side-of-object responses elicited by geometric shapes and natural scenes, it is useful to first evaluate the differences found within various geometric shapes, which were defined using each cell's optimal color and gray.

In Figure 4.3, I plotted the side-of-object effects elicited by the c-shape figure against the side-of-object effect of the standard squares for each cell that was presented with both stimuli (Figure 4.3a). Similarly, the side-of-object effects of the overlapping squares are compared to the standard squares (Figure 4.3b). I fitted the population data using orthogonal regression through the origin and performed bootstrap on the cells in order to calculate the 95% confidence interval. In both cases the slopes were positive, indicating that side-of-object preference was consistent. However, the side-of-object effect elicited by the c-shape figure as about half as strong as the side-of-object effect elicited by standard square: the slope of 0.49 in the population fit for the c-shape figure and standard square (with 95% confidence interval of 0.28 to 0.79). The strength of the border-ownership signal elicited by the overlapping squares was even weaker. The slope of the population fit for the overlapping squares and standard square was 0.29, indicating that the border-ownership signal elicited by

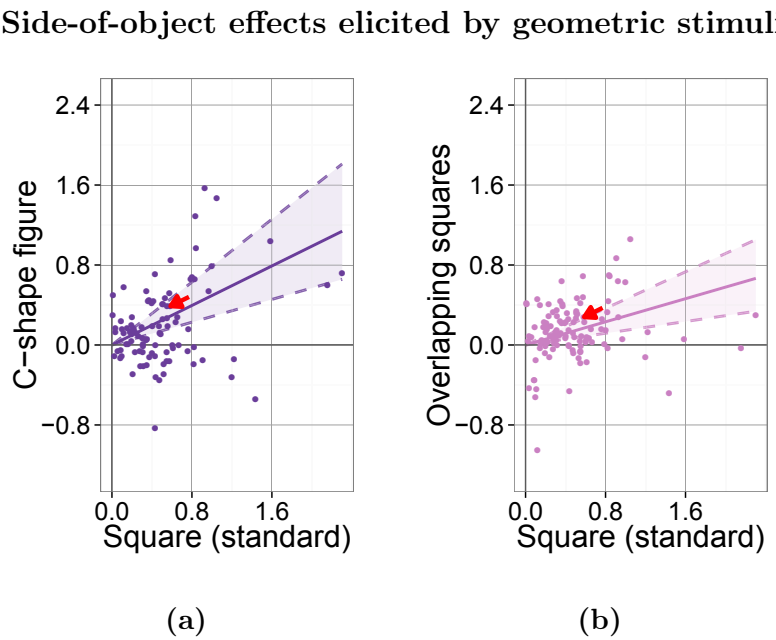


Figure 4.3: Comparison of border-ownership effects of V2 neurons between different geometric stimuli, normalized by the standard test. The side-of-object effects elicited by the standard square test is plotted along the x-axis and side-of-object effects elicited by a) the c-shape and b) the overlapping squares are plotted along the y-axis for all the cells presented with the corresponding stimulus types (96 cells and 113 cells, respectively). The solid lines indicate the population fits and the shading and dashed lines indicate the 95% confidence intervals. The example neuron (28ik1a) is indicated with a red arrow.

the overlapping squares was about a quarter of the strength of the standard square (with 95% confidence interval of 0.15 to 0.46).

Using the modulation index instead of the normalized effect slightly increased the border-ownership signal elicited by the overlapping squares compared to the standard squares (Figure 4.4): the population fit for the overlapping squares and standard square was 0.31, indicating that the border-ownership modulation index elicited by the overlapping squares was about one-third of the strength of the standard square (with 95% confidence interval of 0.19 to 0.49). There was not much change when the modulation index was used for the c-shape figure: the c-shape figure elicited 0.48 the amount of border-ownership modulation compared to the standard square by the modulation index (with 95% confidence interval of 0.29 to 0.69).

4.2.2 Effect of natural color and texture on squares

There was a borderline significant decrease in border-ownership signal strength when the colors (or both the colors and textures) from the occluding object region and background region surrounding each natural scene point were applied to squares and their backgrounds, respectively (Figure 4.5). The slope of the population fit for the color matched squares was 0.83 (with 95% confidence interval of 0.70 to 0.99). Adding the natural-inspired textures reduced the border-ownership signal to 0.76 of that of the standard test (with 95% confidence interval of 0.59 to 1.02). Applying the shapes of the natural objects ("silhouettes") reduced the border-ownership signal to 0.77 of

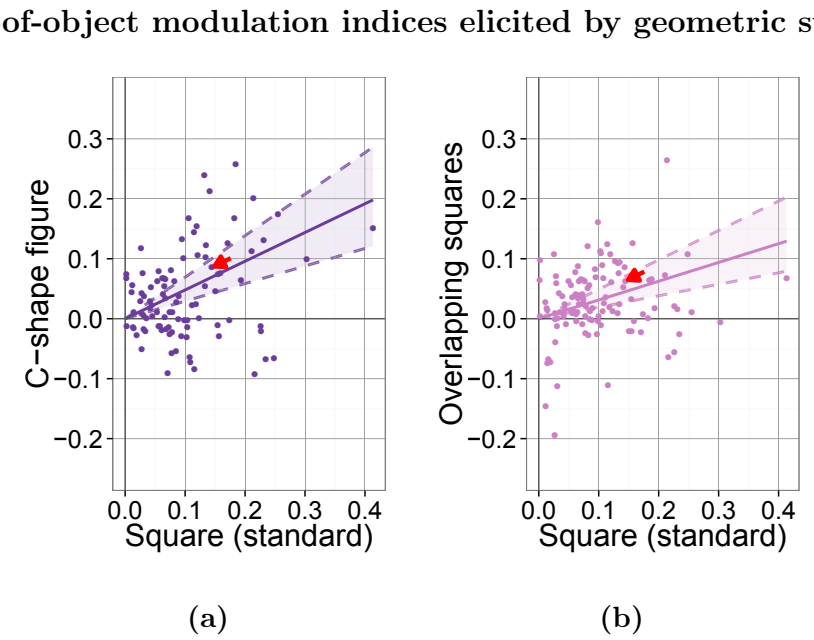


Figure 4.4: Same plot as Figure 4.3, except with the modulation indices plotted instead of the normalized effects. The population fits were quite similar.

that of the standard test (with 95% confidence interval of 0.60 to 0.97).

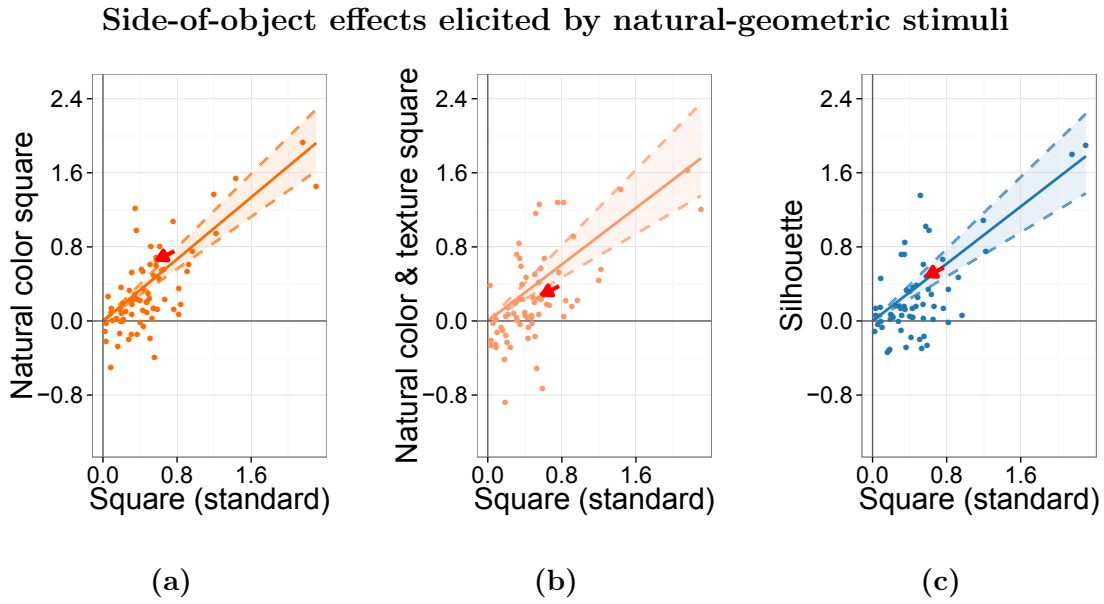


Figure 4.5: Comparison of border-ownership effects of V2 neurons between standard square and hybrid stimuli, normalized by the residuals from the standard test. The border-ownership signal elicited by the standard square test is plotted along the x-axis and border-ownership signal elicited by a) the color matched squares, b) the color and textured matched squares, and c) object silhouettes are plotted along the y-axis for all the cells presented with the corresponding stimulus types (75, 72, and 63 cells, respectively).

These reductions may be due in part because of the reduction of overall firing rates with less preferred stimuli. While there is still a reduction for both cases compared to the standard square when the border-ownership modulation indices were used, these were no longer significant (Figure 4.6 and 4.7). The slope of the population fit for the color matched square was 0.89 (with 95% confidence interval of 0.72 to 1.08).

Adding the natural-inspired texture reduced the border-ownership signal to 0.95 of that of the standard test (with 95% confidence intervals of 0.66 to 1.39). Applying the shape of the natural objects reduced the border-ownership signal to 0.83 of that of the standard test (with 95% confidence interval of 0.62 to 1.15).

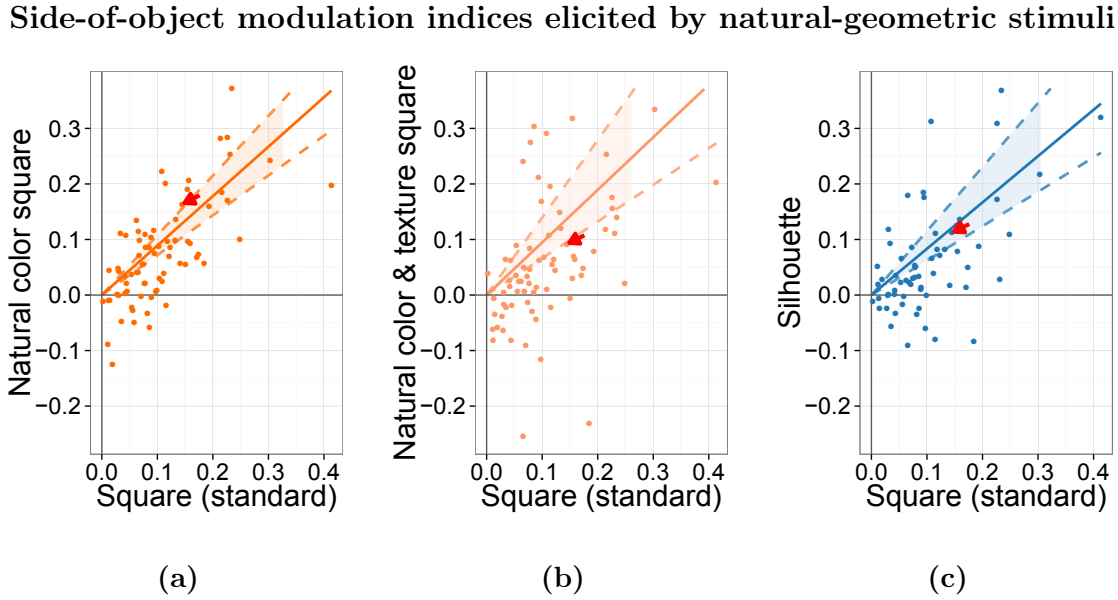
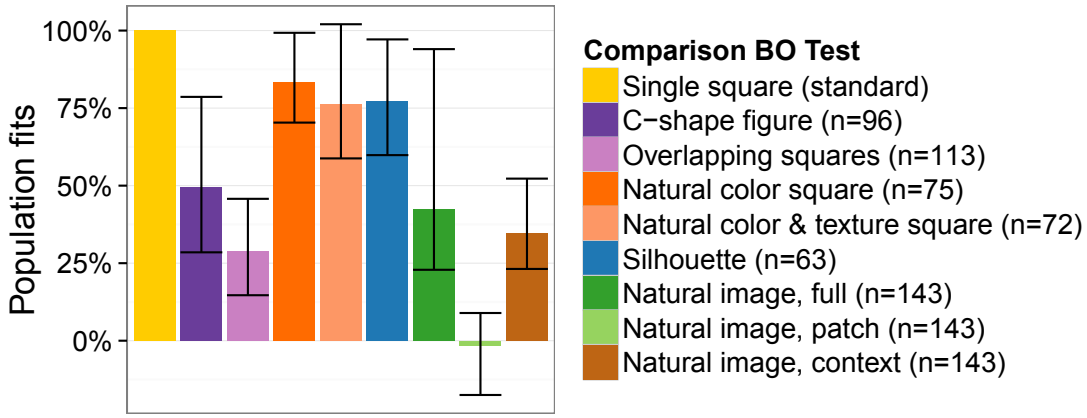
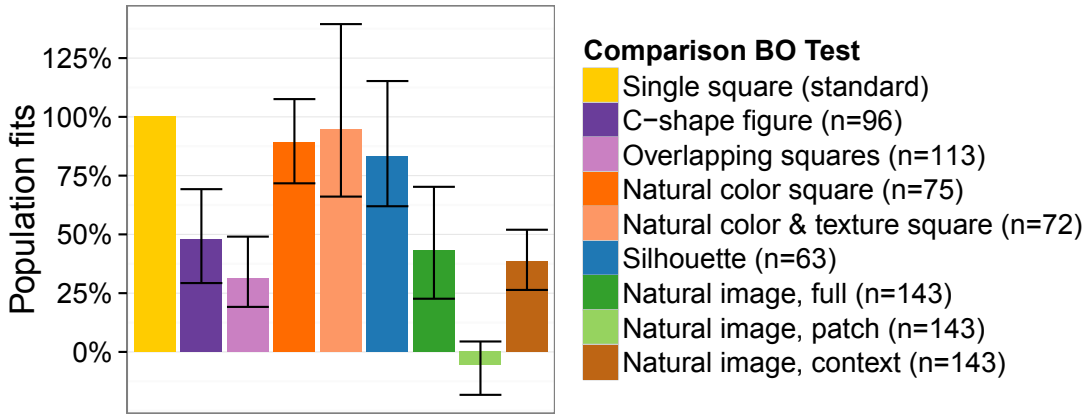


Figure 4.6: Same plot as Figure 4.5, except with the modulation indices plotted instead of the normalized effects.

A more authentic reproduction of textures may have produced different results. While certain frequency components of the textures from the natural scenes were reproduced in the textured squares, a lot of lower frequency components were left out. Extremal edges and shading were also not reproduced in the textures.



(a) Side-of-object normalized effects



(b) Side-of-object modulation index

Figure 4.7: Population fits and confidence intervals using a) normalized effects (summarizing Figures 4.3 (geometric), 4.5 (hybrid), and 3.3 (natural)) and b) modulation indices (summarizing Figures 4.4 (geometric), 4.6 (hybrid), and 3.4 (natural)).

4.2.3 Effect of uniform textures and silhouettes

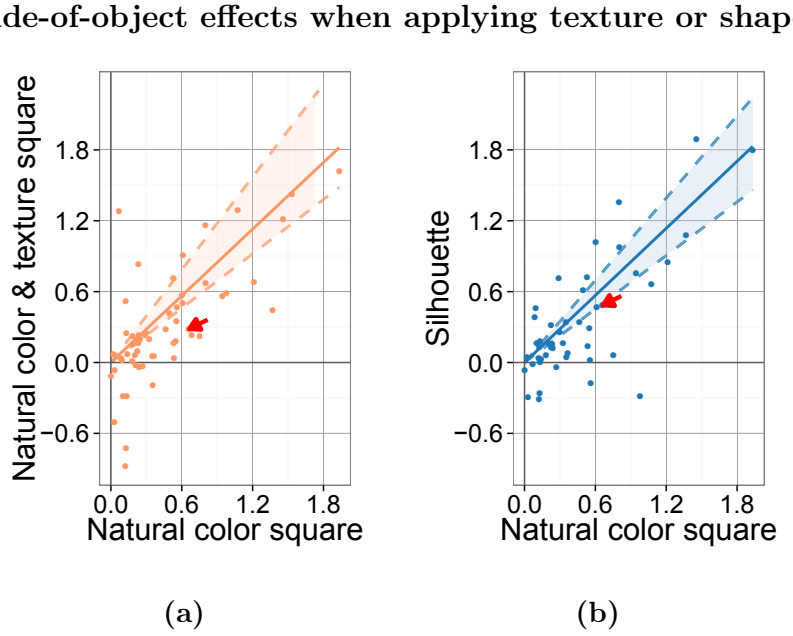


Figure 4.8: Comparison of side-of-object effects of V2 neurons when adding texture and silhouette shape. The side-of-object signal elicited by the color-matched square test is plotted along the x-axis and side-of-object effect elicited by a) the color and texture matched squares and b) the silhouette shapes are plotted along the y-axis for all the cells presented with the corresponding stimulus types (72 cells and 63 cells, respectively). The example neuron (28ik1a) is indicated with a red arrow.

There was no significant effect on the border-ownership signal strength of applying either the texture or object shape on the natural color-matched square (Figure 4.8). Adding the texture had the effect of multiplying the side-of-object effect by 0.94 (95% confidence interval: 0.77 to 1.31). Adding the silhouette had the effect of multiplying

the side-of-object effect by 0.95 (95% confidence interval: 0.76 to 1.16).

4.2.4 Comparison of silhouettes with natural scenes

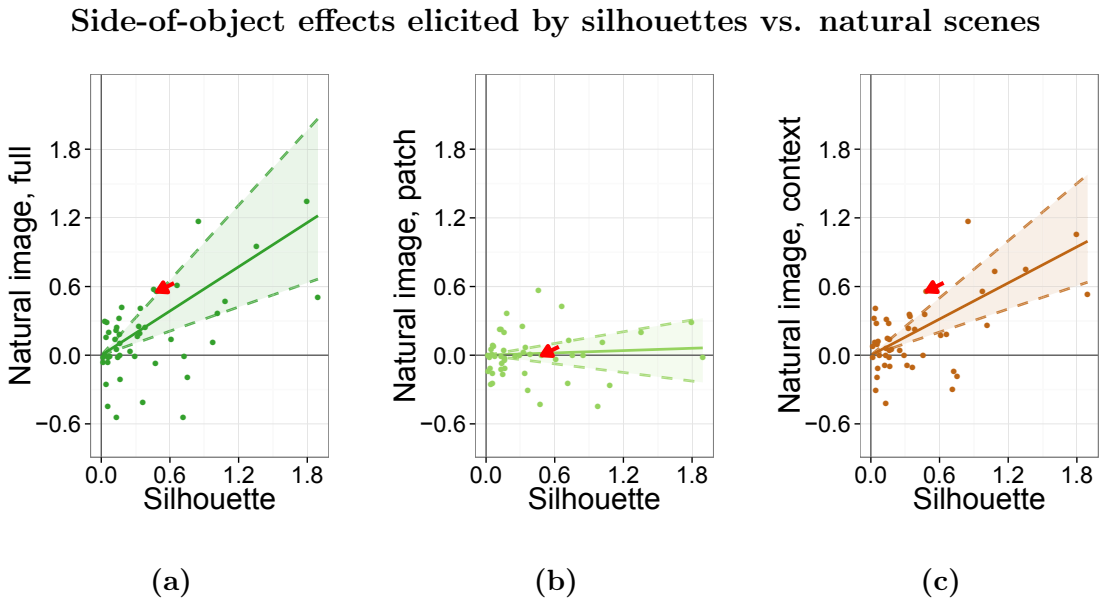


Figure 4.9: V2 population comparison of side-of-object effects of silhouette shapes and natural scenes.

There was a significant reduction in the border-ownership signal strength elicited in the global context of the natural scenes compared to the silhouettes. The border-ownership signal elicited by the global context of the natural scenes was on average 0.53 that of the silhouettes (95% confidence interval: 0.34 to 0.83). The border-ownership signal strength elicited by the full images was weaker relative to the silhouettes, however, the reduction was not significant. The border-ownership signal

elicited by the full images of the natural scenes was on average 0.64 that of the silhouettes (95% confidence interval: 0.35 to 1.09). The border-ownership signal elicited by the local patches of the natural scenes was on average 0.034 that of the silhouettes (95% confidence interval: -0.13 to 0.17).

When the side-of-object modulation indices were used, the population slope fits were similar as when the normalized effects was used. The slopes of the population fits were 0.67 (confidence interval: 0.34 to 1.25) for the full images, 0.02 (confidence interval: -0.16 to 0.24) for the local patches, 0.53 (confidence interval: 0.32 to 0.84) for the global context of the natural scenes when compared to the silhouettes.

Hence, the shape and color of the object is responsible for less of the reduction in the border-ownership signal strength than all of the other complexities included in the natural scenes (such as textures, lighting, and the other objects). This is not surprising since a lot of complexity of the image was not captured in the object silhouettes.

4.3 Conclusion

These results show that border-ownership signal strength decreases with the complexity of the stimulus. This is shown in geometric shapes, with the comparison of the single square with the c-shape figure and overlapping square stimulus types (Figure 4.3). Additionally, the increase in complexity going from silhouettes of objects in natural scenes to the original natural scenes also caused a significant decrease in

border-ownership signal (Figure 4.9).

There was a borderline significant decrease in border-ownership signal when either the natural color was applied to squares or both the natural color and texture of a scene was applied to squares (Figures 4.5 and 4.7). However, this may partially be due to a reduction in firing rate with less preferred color contrasts, as this was only significant when comparing the side-of-object effects and not the side-of-object modulation indices (Figures 4.6 and 4.7).

Chapter 5

Conclusion

5.1 Summary of findings

The use of artificial stimuli is important in neuroscience because it allows the infinite dimensionality of real world experience to be reduced to a manageable number of dimensions of interest. However, this benefit comes with the uncertainty of how well the result extends to real-world conditions and whether neurons are really responding to the phenomena of interest, a correlated phenomenon, or some artifact. Certainly, if the neuronal response studied with artificial stimuli does not generalize in naturally occurring conditions, then the functional significance must be limited (Felsen and Dan 2005).

This thesis demonstrated that the visual area V2 of primates does indeed code for border-ownership in natural scenes. This was done by first demonstrating that a population of V2 cells has a significant effect of side-of-object (46% and 27% of the cells in full images and global context, respectively), i.e., there are some neurons that fire at a higher rate on average when an object appears on one side of the classical

receptive field (CRF) (the cell's preferred side-of-object) than when the object appears on the other side. It was then shown that there is a significant consistency across the V2 population between the side-of-object effects elicited from the global context (but not the local patch) of the natural scene and the standard square test (Figure 3.3). This establishes that the side-of-object effects reflect border-ownership coding in the neural population.

Another interesting finding is that while the border-ownership signal from the global context of the natural scenes appeared later than the border-ownership signal from the single square stimulus, it emerged as early as 75 ms and plateaued by 150 ms (Figure 3.5).

The distribution of the side-of-object effects across the scene points (of a population of border-ownership significant cells) showed that the cells did a remarkable job in coding for the border-ownership, especially once the trial-to-trial variation was taken into account (Figure 3.8). While the local patches of the individual scene points elicited a lot of side-of-object effects that were not just caused by the trial-to-trial variation, there was no consistent agreement between the scene points and there was no significant overall effect. This is indicated by the side-of-object response differences of the local patch of each scene point being approximately equally split between eliciting a more positive and eliciting a more negative response than the null hypothesis line in the Figure 3.8, middle plot. Amazingly, almost all of the side-of-object effects from the global context were in agreement with the standard test, once the variation

expected from trial-to-trial variation was taken into account (global context plot in Figure 3.8).

I failed to find a significant positive correlation of the overall surround suppression and the border-ownership coding, in fact, there was a slightly negative correlation (Figure 3.9).

The potential concern that rotating the images upside down would significantly reduce the neural response was addressed by plotting the effect of being right side up versus upside-down in Figure 3.10, which showed that there was no such effect in the population.

I also found that the border-ownership signal from natural scenes is significantly weaker than that of the standard square test (Figure 3.3). By presenting additional geometric stimuli and stimuli that were hybrids of geometric figures and natural scenes, I was able to show some of the dimensions that correlate with a reduction in border-ownership coding. Both the c-shape and overlapping squares elicited a weaker border-ownership signal than the standard square test (Figure 4.3). Applying the natural color (with or without the texture) to the squares instead of each cell's preferred color resulted in a borderline significant reduction in the elicited border-ownership signal (Figure 4.5a-b). Using the object's shape, in addition to natural color, resulted in a non-significant reduction in the border-ownership signal (Figure 4.5c). There was also a significant reduction in the border-ownership signal strength elicited in the global context of the natural scenes compared to the silhouettes, that

is, silhouettes of the objects elicited a much stronger signal than the objects in the natural scene context (Figure 4.9).

5.2 Discussion

The ventral stream transforms neural responses from being retinotopic to being object-centered and largely invariant to position (Connor, Brincat, and Pasupathy 2007). For example, the posterior inferotemporal cortex (pIT/TEO) contains cells that seem to respond to combinations of curvatures in relationship to the center of the object (Brincat and Connor 2004) and central and anterior inferotemporal cortex contain cells that encode for three-dimensional shapes (Yamane et al. 2008). However, the task of perceptually organizing visual regions into objects is extremely difficult in natural scenes and remains a poorly solved problem in computer vision. It is unknown how the visual system is able to accomplish this challenging task.

This thesis work is the first neurophysiological study of border-ownership coding in the early visual cortex using natural scenes. The following subsections give context to the findings in this thesis and discuss the broader implications for perceptual grouping and natural vision processing.

5.2.1 Contribution to understanding border-ownership coding and perceptual grouping

One of the earliest indications of perceptual grouping to occur after the onset of a new image is the phenomenon of border-ownership coding in the early visual cortex, where neurons exhibit a higher firing rate when an occluding figure is on one side of their receptive field, the preferred side, versus the other. This is observable as early as the supragranular layers of V1 and commonly appears in V2 and V4. Learning that border-ownership coding does occur for natural scenes provides further evidence that border-ownership plays a significant role in perceptual grouping. The delay in the border-ownership signals for natural scenes also provides further evidence that this process is not a feed-forward process. On the other hand, the reduction in strength of the border-ownership signal in natural scenes and the speed in which the border-ownership signal plateaus (within 150 ms) provides evidence that this signal is not due to feedback from very high-level visual areas. In other words, this early border-ownership coding likely plays a significant role in the determination of the final border-ownership and perceptual grouping of a scene, but does not seem to reflect this final determination.

There are many local cues that influence our perception of border-ownership (see Section 1.1). Surprisingly, the CRFs did not contribute significantly to the border-ownership signal (a population side-of-object signal that correlates with the standard

square test). Although there were more cells with a significant effect of side-of-object from the local patches than what would be expected from chance, these signals did not significantly correlate with the standard square test. And while it is possible that there still may be a locally derived component of the border-ownership signal, it appears that this contribution would be very small (less than 11% of the signal strength of the standard test, $p < 0.025$). This is surprising because the statistics of natural scenes allow for border-ownership to be determined from the local context well above chance (Ramenahalli, Mihalas, and Niebur 2014). While these cells contain information that is known to influence the perception of border-ownership, it does not seem to be reflected in this early border-ownership signal. This does not mean that this local information is not used in the higher-levels of vision, but it appears that this information is not integrated into the border-ownership signal at a low level.

Why do natural scenes elicit weaker border-ownership modulation than single squares? A square can also be interpreted as a "window" or a square hole in a flat surface. Natural scenes have a lot more cues that can be used to assign border-ownership, such that it is practically impossible to interpret border-ownership in any other way. It is clear from the results of this thesis that the strength of the border-ownership modulation is not representing the probability of border-ownership. If this was the case, the standard square test would have weaker border-ownership modulation compared to the overlapping square test and the full images of the natural scenes, since the overlapping squares and natural scenes have more cues for border-

ownership assignment (Section 1.1).

Our lab's model of border-ownership coding (Figure 1.12C), however, predicts that complex natural scenes would elicit weaker and less reliable border-ownership coding than isolated squares. In this model, a strong border-ownership modulation would be calculated when there are many edges and corners (or other textures) on the side of the object and no edges or corners on the side with no object (with just a solid background), such is the case with our standard square test. However, with natural scenes there are often many edges and texture on both the side of the object and the background, hence there would be a reduction in the border-ownership modulation strength.

Hence, the border-ownership modulations in the primate visual area V2 are not calculated from very high levels of the brain that have aggregated all of the global information from a scene and would be able to provide a probability of border-ownership assignments given all of the local and global cues. Instead, these modulations are the result of quick heuristics that have limited access to high level object information.

5.2.2 Contribution to understanding visual processing of natural images

Carandini et al. (2005) have stated, "We can claim that we know what the visual system does once we can predict neural responses to arbitrary stimuli, including

those seen in nature." While it is often assumed that the early visual system is well characterized, we are far from being able to fulfill this criterion. Current models of V1 and V2 neurons can only account for 30-40% of the explainable variance of the responses to new stimuli (Carandini et al. 2005; Olshausen and Field 2005; Willmore, Prenger, and Gallant 2010), even under the fixation paradigm using gray-scale natural images in which context is limited by presenting large patches.

Since the main goal of this thesis research has been to study border-ownership and not to build a comprehensive model of visual neurons, I did not present a sufficient number of images to be able to determine whether accounting for border-ownership can improve the amount of explainable variance or to use their methods for modeling the receptive field.

David, Vinje, and Gallant (2004) showed that there is a slow temporal inhibitory component in responses to natural stimuli that did not appear with artificial gratings. The same lab later extended their methods to analyze V2 neurons and found that there were two main clusters of V2 neurons. Cluster A neurons followed the majority of the responses of V1 and had little to no suppressive tuning, while cluster B neurons had a substantial suppressive tuning (Willmore, Prenger, and Gallant 2010). I expected that neurons with more suppressive tuning (type B) would tend to have more effect of border-ownership. Surprisingly, there was no positive correlation between the suppression modulation and the side-of-object modulation across the population of V2 neurons. This may be because border-ownership cells may rely on a combination

of inhibitory and excitatory regions and this signal may thus be uncorrelated with the overall surround suppression. The models from both Sakai and Nishimura (2006) and Craft et al. (2007) predict that excitatory and inhibitory regions of the extraclassical receptive fields contribute to the calculation of the border-ownership signals. In fact, Zhang and von der Heydt (2010) have found that border-ownership cells do rely on such a combination of excitatory and inhibitory regions.

Further work should evaluate the relationship between these excitatory and inhibitory regions (anisotropic surrounds) and the border-ownership coding in natural scenes. This could be done by using the rapid presentation of natural scenes as in Willmore, Prenger, and Gallant (2010) and seeing how much (if any) these surrounds account for border-ownership coding. The calculated spatiotemporal receptive fields and the clustering would allow for the correlation of the subtypes of V2 neurons and border-ownership to be tested, which would help determine the significance of these two subtypes. Additionally, such an experiment could give us a better understanding of the mechanisms of how border-ownership is calculated by determining the contribution of the extraclassical receptive fields to border-ownership coding. If the spatiotemporal receptive fields calculated by the rapid natural scene presentation does not help explain the border-ownership, this could give us insight that would help to improve this technique so that it could elucidate other findings that this method may have missed.

Using stationary natural scenes with the fixation paradigm is only a step towards

studying the visual system in natural conditions. Future work should also allow for free eye movements, where the subjects are not required to maintain fixation. This is important because staring at a single spot for several seconds is very unnatural and the subjects have to actively suppress eye movements. In natural conditions, primates make eye movements several times a second, and yet the visual system is able to maintain a stable perception of the world. Additionally, natural scenes should be allowed to be dynamic, either by the subject's movements or by the movement of objects in the scene.

Appendix A

Border-ownership coding of natural scenes in primate V1

A total of 38 V1 cells were tested for an effect of side-of-object using both the standard square and the natural scene tests (with two or more presentations of at least 10 scene points). The majority of these cells are from one hemisphere from Monkey M27 (see Table 2.1). Of these 38 cells, 7 cells (18%) had a significant ($p < 0.01$) side-of-square effect, 15 cells (39%) had a a significant effect of side-of-object with the full images, 14 cells (37%) had a a significant effect of side-of-object with the local patches, and 3 cells (8%) had a a significant effect of side-of-object with the global context (decomposed from full and patch responses, see Section 2.8.1).

The responses of the V1 neurons are shown in Figure A.1. The slope of the population fit for the natural scenes was 1.48 (with 95% confidence interval of 0.01 to 1021.45) for the full images, -0.1 (with 95% confidence interval of -7.74 to 1.87) for the local patches, and 0.18 (with 95% confidence interval of -0.04 to 0.3) for the global context.

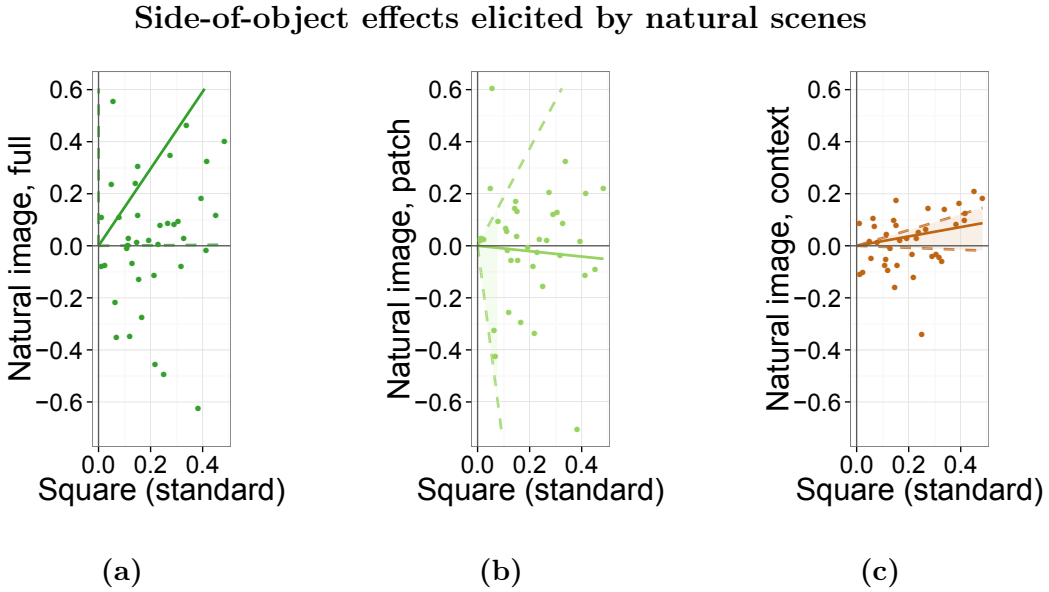


Figure A.1: V1 population comparison of side-of-object effects between standard square and natural stimuli. The side-of-object effect elicited by the standard square test is plotted along the x-axis and side-of-object effect elicited by a) full images, b) local classical receptive field (CRF) patches, and c) global context are plotted along the y-axis for all the V1 cells presented with at least 10 scene points (with all eight variations presented at least twice, 38 cells).

Appendix B

Comparison across monkeys

In this chapter, I split up the results from visual area V2 by the three monkeys. The comparison of the effect of side-of-object from the natural images to the standard test is shown in Figures B.1, B.2, and B.3. The comparison of the effect of side-of-object from the geometric stimuli is shown in Figures B.4, B.5, and B.6. The comparison of the effect of side-of-object from the natural-geometric "hybrid" stimuli is shown in Figures B.7 and B.8.

Monkey M27 only had three V2 cells that were presented with natural color and natural textured squares, and was not presented any silhouette test. All of these cells elicited approximately the effect of side-of-object between all three square tests (standard/optimal color, natural color, and natural color and texture; not plotted).

The results are summarized in Figure B.9. A major difference between the monkeys is that there was a marked reduction in the c-shape figure and overlapping square tests in the neurons from monkey M29.

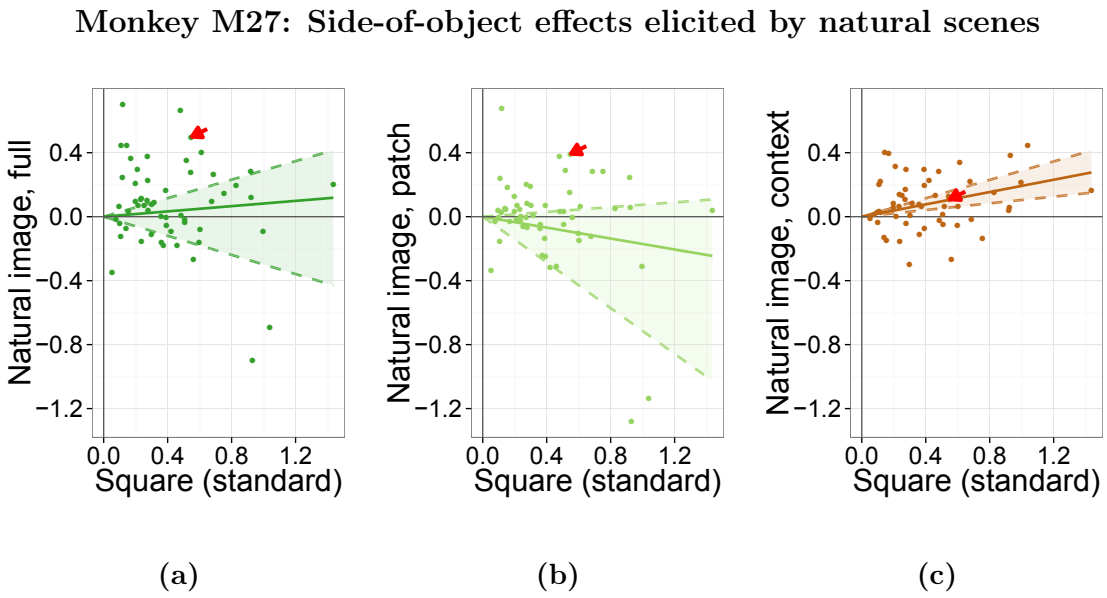


Figure B.1: Monkey M27: Population comparison of side-of-object effects between standard square and natural stimuli. The side-of-object effect elicited by the standard square test is plotted along the x-axis and side-of-object effect elicited by a) full images, b) local classical receptive field (CRF) patches, and c) global context are plotted along the y-axis for all the V2 cells from monkey M27 presented with at least 10 scene points (with all eight variations presented at least twice, 56 cells). The example neuron 27lj2d is indicated with a red arrow.

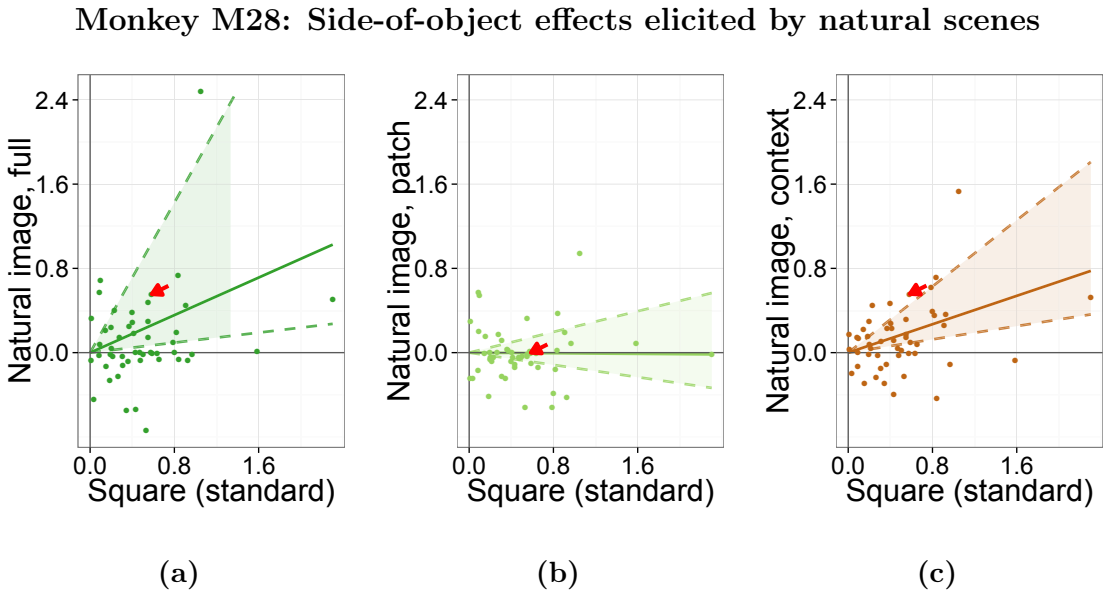


Figure B.2: Monkey M28: Population comparison of side-of-object effects between standard square and natural stimuli. The side-of-object effect elicited by the standard square test is plotted along the x-axis and side-of-object effect elicited by a) full images, b) local CRF patches, and c) global context are plotted along the y-axis for all the V2 cells from monkey M28 presented with at least 10 scene points (with all eight variations presented at least twice, 50 cells). The example neuron 28ik1a is indicated with a red arrow.

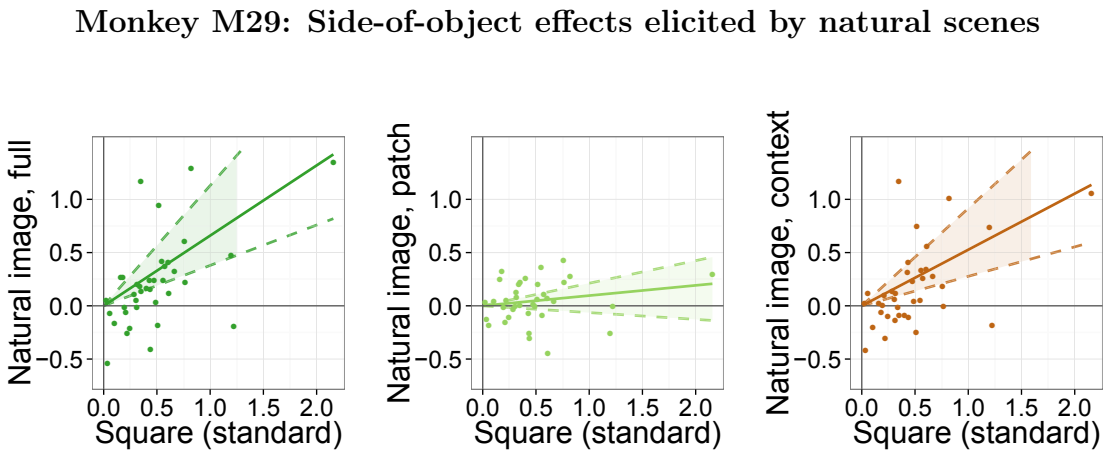
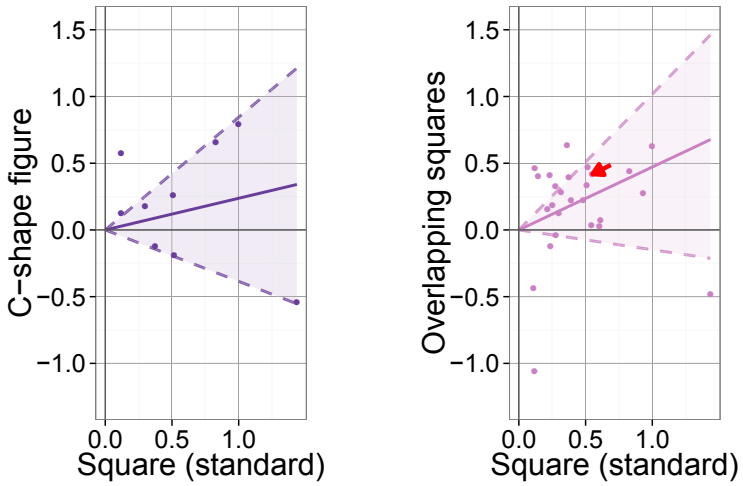


Figure B.3: Monkey M29: Population comparison of side-of-object effects between standard square and natural stimuli. The side-of-object effect elicited by the standard square test is plotted along the x-axis and side-of-object effect elicited by a) full images, b) local CRF patches, and c) global context are plotted along the y-axis for all the V2 cells from monkey M29 presented with at least 10 scene points (with all eight variations presented at least twice, 37 cells).

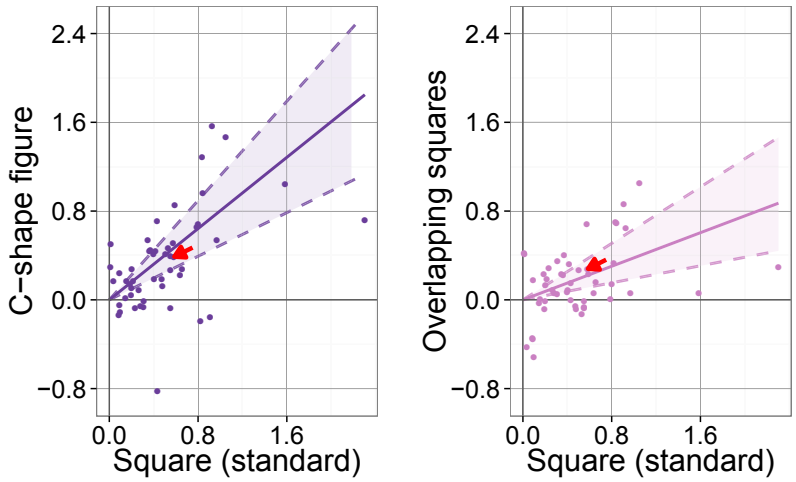
Monkey M27: Side-of-object effects elicited by geometric stimuli



(a) **(b)**

Figure B.4: Monkey M27: Comparison of border-ownership effects of V2 neurons between different geometric stimuli, normalized by the standard square test. The side-of-object effects elicited by the standard square test is plotted along the x-axis and side-of-object effects elicited by a) the c-shape and b) the overlapping squares are plotted along the y-axis for all the cells presented with the corresponding stimulus types (9 cells and 26 cells, respectively). The 95% confidence interval is indicated by the dashed lines. The example neuron 27lj2d is indicated with a red arrow for overlapping squares test (cell not presented with c-shape figure test).

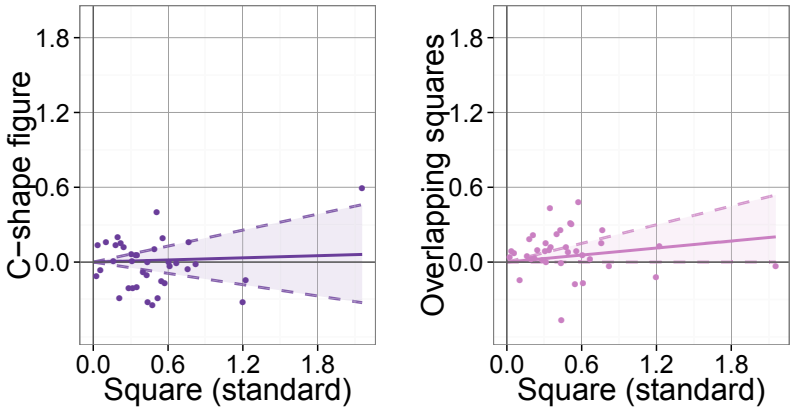
Monkey M28: Side-of-object effects elicited by geometric stimuli



(a) **(b)**

Figure B.5: Monkey M28: Comparison of border-ownership effects of V2 neurons between different geometric stimuli, normalized by the standard test. The side-of-object effects elicited by the standard square test is plotted along the x-axis and side-of-object effects elicited by a) the c-shape and b) the overlapping squares are plotted along the y-axis for all the cells presented with the corresponding stimulus types (50 cells and 50 cells, respectively). The 95% confidence interval is indicated by the dashed lines. The example neuron 28ik1a is indicated with a red arrow.

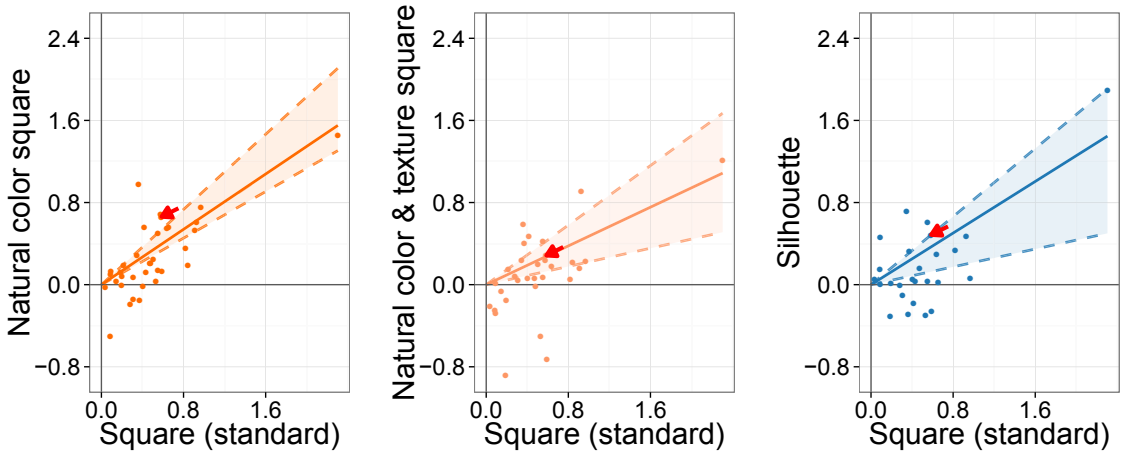
Monkey M29: Side-of-object effects elicited by geometric stimuli



(a) **(b)**

Figure B.6: Monkey M29: Comparison of border-ownership effects of V2 neurons between different geometric stimuli, normalized by the standard test. The side-of-object effects elicited by the standard square test is plotted along the x-axis and side-of-object effects elicited by a) the c-shape and b) the overlapping squares are plotted along the y-axis for all the cells presented with the corresponding stimulus types (37 cells and 37 cells, respectively). The 95% confidence interval is indicated by the dashed lines.

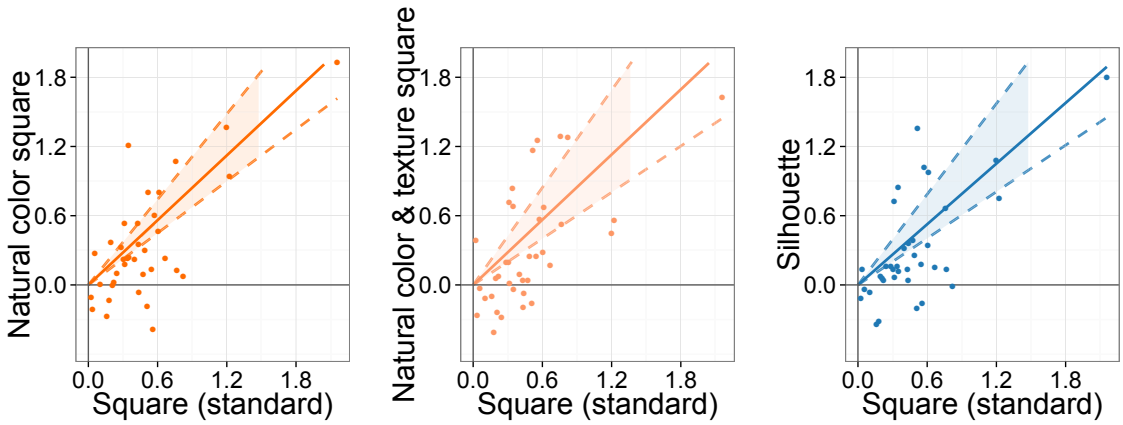
Monkey M28: Side-of-object effects elicited by natural-geometric stimuli



(a) **(b)** **(c)**

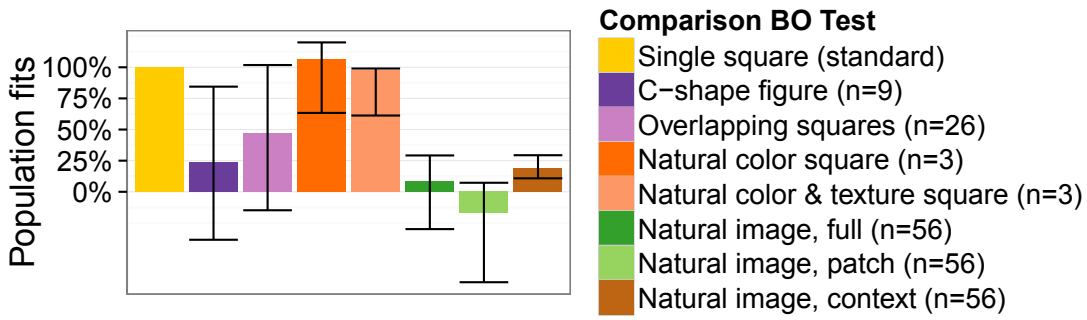
Figure B.7: Monkey 28: Comparison of border-ownership effects of V2 neurons between standard square and hybrid stimuli, normalized by the residuals from the standard test. The border-ownership signal elicited by the standard square test is plotted along the x-axis and border-ownership signal elicited by a) the color matched squares, b) the color and textured matched squares, and c) object silhouettes are plotted along the y-axis for all the cells presented with the corresponding stimulus types (35, 32, and 26 cells, respectively).

Monkey M29: Side-of-object effects elicited by natural-geometric stimuli

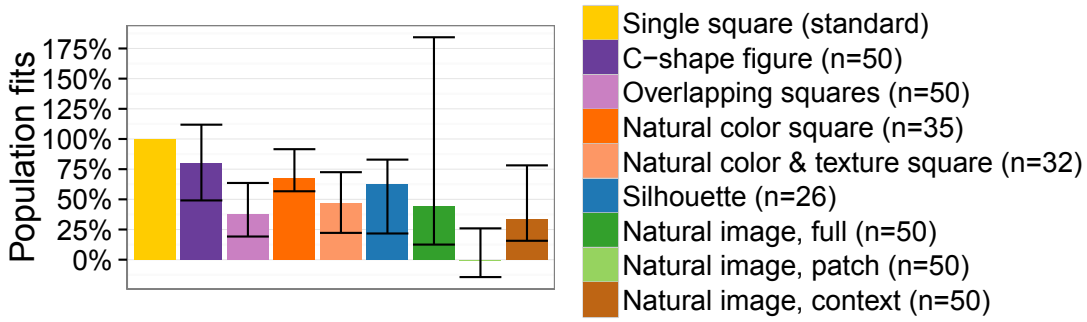


(a) **(b)** **(c)**

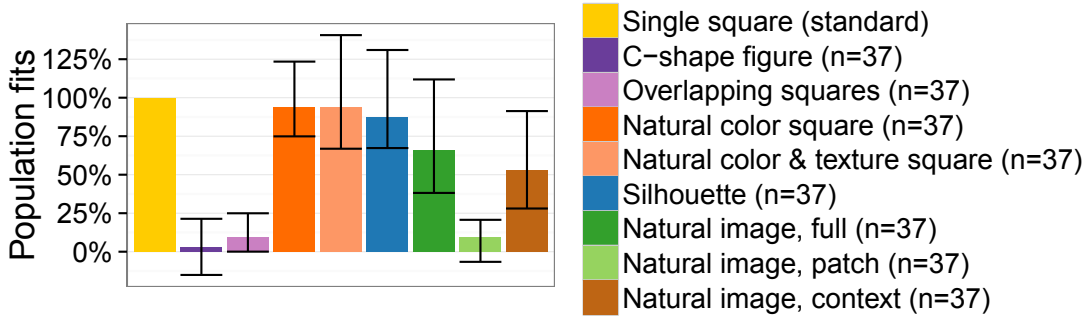
Figure B.8: Monkey 29: Comparison of border-ownership effects of V2 neurons between standard square and hybrid stimuli, normalized by the residuals from the standard test. The border-ownership signal elicited by the standard square test is plotted along the x-axis and border-ownership signal elicited by a) the color matched squares, b) the color and textured matched squares, and c) object silhouettes are plotted along the y-axis for all the cells presented with the corresponding stimulus types (37, 37, and 37 cells, respectively).



(a) M27: side-of-object normalized effects



(b) M28: side-of-object normalized effects



(c) M29: side-of-object normalized effects

Figure B.9: Summary of comparison of border-ownership effects elicited by different stimuli across monkeys.

Glossary

ANOVA analysis of variance.

border-ownership A border that is created by visual occlusion is said to be owned by the nearer, occluding object.

border-ownership coding The coding of border-ownership in the early visual cortex via side-of-object modulations.

BSDS300 Berkeley Segmentation Dataset 300.

CRF classical receptive field.

pIT posterior inferior temporal cortex.

PSTH post-stimulus histogram.

r_{crf^+} The radius in which the stimulus within the patch is exactly the same as in the global context. Equal or slightly larger than the radius of the classical receptive field. See Figure 2.3.

scene point A point or position in a specific scene. In my experiments, scene points always lie on occlusion borders.

side-of-object The location of an object relative to a classical receptive field (CRF). It can be on one of two sides, orthogonal to the preferred orientation of the neuron.

side-of-object modulation A modulation in a neuron's firing rate due to the side-of-object.

Bibliography

Albright, T. D. and Stoner, G. R. (2002). "Contextual Influences on Visual Processing". In: *Annual Review of Neuroscience* 25.1, pp. 339–379. DOI: 10.1146/annurev.neuro.25.112701.142900. URL: http://dx.doi.org/10.1146/annurev.neuro.25.112701.142900.

Angelucci, A., Levitt, J. B., and Lund, J. S. (2002). "Anatomical origins of the classical receptive field and modulatory surround field of single neurons in macaque visual cortical area V1". In: *Progress in Brain Research* 136, pp. 373–388.

Balleis, J. (2013). *Xiao Liwu im San Diego Zoo*. URL: http://commons.wikimedia.org/wiki/File:Xiao_Liwu_im_San_Diego_Zoo_-_Foto_3.jpeg.

Bregman, A. S. (1981). "Asking the "what for" question in auditory perception". In: *Perceptual organization*. Ed. by M. Kubovy and J. R. Pomerantz, pp. 99–118.

Brincat, S. L. and Connor, C. E. (2004). "Underlying principles of visual shape selectivity in posterior inferotemporal cortex". In: *Nature Neuroscience* 7.8, pp. 880–886. DOI: 10.1038/nn1278. URL: http://www.nature.com/neuro/journal/v7/n8/full/nn1278.html.

Canty, A. and Ripley, B. D. (2014). *boot: Bootstrap R (S-Plus) Functions*. R package version 1.3-11.

BIBLIOGRAPHY

Carandini, M. et al. (2005). "Do We Know What the Early Visual System Does?" In: *The Journal of Neuroscience* 25.46, pp. 10577–10597. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.3726-05.2005. URL: http://www.jneurosci.org/content/25/46/10577.

Connor, C. E., Brincat, S. L., and Pasupathy, A. (2007). "Transformation of shape information in the ventral pathway". In: *Current Opinion in Neurobiology* 17.2, pp. 140–147.

Craft, E. et al. (2007). "A neural model of figure-ground organization". In: *Journal of Neurophysiology* 97.6, pp. 4310–4326.

David, S. V., Vinje, W. E., and Gallant, J. L. (2004). "Natural Stimulus Statistics Alter the Receptive Field Structure of V1 Neurons". In: *The Journal of Neuroscience* 24.31, pp. 6991–7006. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.1422-04.2004. URL: http://www.jneurosci.org/content/24/31/6991.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications*. ISBN 0-521-57391-2. Cambridge: Cambridge University Press. URL: http://statwww.epfl.ch/davison/BMA/.

Egly, R., Driver, J., and Rafal, R. D. (1994). "Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects." In: *Journal of Experimental Psychology: General* 123.2, p. 161. URL: http://psycnet.apa.org/journals/xge/123/2/161/.

BIBLIOGRAPHY

Felsen, G. and Dan, Y. (2005). "A natural approach to studying vision". In: *Nature Neuroscience* 8.12, pp. 1643–1646. ISSN: 1097-6256. DOI: 10.1038/nn1608. URL: http://www.nature.com/neuro/journal/v8/n12/full/nn1608.html.

Fox, J. (2003). "Effect Displays in R for Generalised Linear Models". In: *Journal of Statistical Software* 8.15, pp. 1–27. URL: http://www.jstatsoft.org/v08/i15/.

Fox, J. and Weisberg, S. (2011). *An R Companion to Applied Regression*. Second ed. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

Girard, P., Hupe, J. M., and Bullier, J. (2001). "Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities". In: *Journal of Neurophysiology* 85.3, pp. 1328–1331. URL: http://jn.physiology.org/content/85/3/1328.short.

Good, W. V. et al. (2001). "Recent advances in cortical visual impairment". In: *Developmental Medicine & Child Neurology* 43.1, pp. 56–60. ISSN: 1469-8749. DOI: 10.1111/j.1469-8749.2001.tb00387.x. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.1469-8749.2001.tb00387.x/abstract.

Jehee, J. F. M., Lamme, V. A. F., and Roelfsema, P. R. (2007). "Boundary assignment in a recurrent network architecture". In: *Vision Research* 47.9, pp. 1153–1165. ISSN: 0042-6989. DOI: 10.1016/j.visres.2006.12.018. URL: http://www.sciencedirect.com/science/article/pii/S0042698906005700.

BIBLIOGRAPHY

Klymenko, V. and Weisstein, N. (1986). "Spatial frequency differences can determine figure-ground organization". In: *Journal of Experimental Psychology: Human Perception and Performance* 12.3, pp. 324–330. ISSN: 1939-1277(Electronic);0096-1523(Print). DOI: 10.1037/0096-1523.12.3.324.

Lamme, V. A. (1995). "The neurophysiology of figure-ground segregation in primary visual cortex". In: *The Journal of Neuroscience* 15.2, pp. 1605–1615. URL: http://www.jneurosci.org/content/15/2/1605.short.

Marshall, J. A. et al. (1996). "Occlusion edge blur: a cue to relative visual depth". In: *Journal of the Optical Society of America A* 13.4, p. 681. ISSN: 1084-7529, 1520-8532. DOI: 10.1364/JOSAA.13.000681. URL: http://www.opticsinfobase.org/josaa/fulltext.cfm?uri=josaa-13-4-681&id=17456.

Martin, D. et al. (2001). "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics". In: *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings*. Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings. Vol. 2, 416–423 vol.2. DOI: 10.1109/ICCV.2001.937655.

Mihalas, S. et al. (2011). "Mechanisms of perceptual organization provide auto-zoom and auto-localization for attention to objects". In: *Proceedings of the National Academy of Sciences* 108.18, pp. 7583–7588. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1014655108. URL: http://www.pnas.org/content/108/18/7583.

BIBLIOGRAPHY

Neo, P. *Red apple from top*. URL: http://www.public-domain-image.com/flora-plants-public-domain-images-pictures/fruits-public-domain-images-pictures/apple-pictures/red-apple-from-top.jpg.html.

O'Herron, P. and von der Heydt, R. (2009). "Short-term memory for figure-ground organization in the visual cortex". In: *Neuron* 61.5, pp. 801–809.

O'Herron, P. and von der Heydt, R. (2011). "Representation of object continuity in the visual cortex". In: *Journal of Vision* 11.2.

O'Herron, P. and von der Heydt, R. (2013). "Remapping of Border Ownership in the Visual Cortex". en. In: *The Journal of Neuroscience* 33.5, pp. 1964–1974. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.2797-12.2013. URL: http://www.jneurosci.org/content/33/5/1964.

Olshausen, B. A. and Field, D. J. (2005). "How Close Are We to Understanding V1?" In: *Neural Computation* 17.8, pp. 1665–1699. ISSN: 0899-7667. DOI: 10.1162/0899766054026639. URL: http://dx.doi.org/10.1162/0899766054026639.

O'Shea, R. P., Blackburn, S. G., and Ono, H. (1994). "Contrast as a depth cue". In: *Vision Research* 34.12, pp. 1595–1604. ISSN: 0042-6989. DOI: 10.1016/0042-6989(94)90116-3. URL: http://www.sciencedirect.com/science/article/pii/0042698994901163.

Palmer, S. E. and Ghose, T. (2008). "Extremal Edge A Powerful Cue to Depth Perception and Figure-Ground Organization". In: *Psychological Science* 19.1, pp. 77–

83. ISSN: 0956-7976, 1467-9280. DOI: 10.1111/j.1467-9280.2008.02049.x. URL: http://pss.sagepub.com/content/19/1/77.

Peterson, M. and Salvagio, E. (2010). "Figure-ground perception". In: *Scholarpedia* 5.4, p. 4320. ISSN: 1941-6016. DOI: 10.4249/scholarpedia.4320. URL: http://www.scholarpedia.org/article/Figure-ground_perception.

Qiu, F. T., Sugihara, T., and von der Heydt, R. (2007). "Figure-ground mechanisms provide structure for selective attention". In: *Nature Neuroscience* 10.11, pp. 1492–1499. ISSN: 10976256.

Qiu, F. T. and von der Heydt, R. (2005). "Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules". In: *Neuron* 47.1, pp. 155–166.

Qiu, F. T. and von der Heydt, R. (2007). "Neural representation of transparent overlay". In: *Nature Neuroscience* 10.3, pp. 283–284.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org/.

Ramenahalli, S., Mihalas, S., and Niebur, E. (2014). "Local spectral anisotropy is a valid cue for figure-ground organization in natural scenes". In: *Vision Research* 103, pp. 116–126. ISSN: 0042-6989. DOI: 10.1016/j.visres.2014.08.012. URL: http://www.sciencedirect.com/science/article/pii/S0042698914001916.

BIBLIOGRAPHY

Rieke, F. and Baylor, D. A. (1998). "Single-photon detection by rod cells of the retina". In: *Reviews of Modern Physics* 70.3, pp. 1027–1036. DOI: 10.1103/RevModPhys.70.1027. URL: http://link.aps.org/doi/10.1103/RevModPhys.70.1027.

Rubin, E. (1915). *Synsoplevede figurer: studier i psykologisk analyse. 1. del*. Gyldendalske Boghandel, Nordisk Forlag.

Sakai, K. and Nishimura, H. (2006). "Surrounding Suppression and Facilitation in the Determination of Border Ownership". In: *Journal of Cognitive Neuroscience* 18.4, pp. 562–579. ISSN: 0898-929X. DOI: 10.1162/jocn.2006.18.4.562. URL: http://dx.doi.org/10.1162/jocn.2006.18.4.562.

Sugihara, T., Qiu, F. T., and von der Heydt, R. (2011). "The speed of context integration in the visual cortex". en. In: *Journal of Neurophysiology* 106.1, pp. 374–385. ISSN: 0022-3077, 1522-1598. DOI: 10.1152/jn.00928.2010. URL: http://jn.physiology.org/content/106/1/374.

Supér, H., Romeo, A., and Keil, M. (2010). "Feed-Forward Segmentation of Figure-Ground and Assignment of Border-Ownership". In: *PLoS ONE* 5.5, e10705. DOI: 10.1371/journal.pone.0010705. URL: http://dx.doi.org/10.1371/journal.pone.0010705.

Troscianko, T. et al. (1991). "The role of colour as a monocular depth cue". In: *Vision Research* 31.11, pp. 1923–1929. ISSN: 0042-6989. DOI: 10.1016/0042-6989(91)90187-A. URL: http://www.sciencedirect.com/science/article/pii/004269899190187A.

BIBLIOGRAPHY

Vecera, S. P., Vogel, E. K., and Woodman, G. F. (2002). "Lower region: a new cue for figure-ground assignment." In: *Journal of Experimental Psychology: General* 131.2, p. 194. URL: http://psycnet.apa.org/psycinfo/2002-13144-003.

von der Heydt, R., Qiu, F. T., and He, Z. J. (2003). "Neural mechanisms in border ownership assignment: motion parallax and gestalt cues". en. In: *Journal of Vision* 3.9, pp. 666–666. ISSN: , 1534-7362. DOI: 10.1167/3.9.666.

von der Heydt, R., Zhou, H., and Friedman, H. S. (2000). "Representation of stereoscopic edges in monkey visual cortex". In: *Vision Research* 40.15, pp. 1955–1967.

Walker, G. A., Ohzawa, I., and Freeman, R. D. (1999). "Asymmetric Suppression Outside the Classical Receptive Field of the Visual Cortex". en. In: *The Journal of Neuroscience* 19.23. PMID: 10575050, pp. 10536–10553. ISSN: 0270-6474, 1529-2401. URL: http://www.jneurosci.org/content/19/23/10536.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York. ISBN: 978-0-387-98140-6. URL: http://had.co.nz/ggplot2/book.

Williford, J. R. and von der Heydt, R. (2013). "Border-ownership coding". In: *Scholarpedia* 8.10, p. 30040. URL: http://scholarpedia.org/article/Border-ownership_coding.

Willmore, B. D. B., Prenger, R. J., and Gallant, J. L. (2010). "Neural Representation of Natural Images in Visual Area V2". In: *The Journal of Neuroscience* 30.6, pp. 2102–2114. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.4099-09.2010. URL: http://www.jneurosci.org/content/30/6/2102.

BIBLIOGRAPHY

Xie, Y. (2013). *Dynamic Documents with R and knitr*. ISBN 978-1482203530. Boca Raton, Florida: Chapman and Hall/CRC. URL: http://yihui.name/knitr/.

Xie, Y. (2014). *knitr: A general-purpose package for dynamic report generation in R*. R package version 1.7. URL: http://yihui.name/knitr/.

Yamane, Y. et al. (2008). "A neural code for three-dimensional object shape in macaque inferotemporal cortex". In: *Nature Neuroscience* 11.11, pp. 1352–1360. ISSN: 1097-6256. DOI: 10.1038/nn.2202. URL: http://www.nature.com/neuro/journal/v11/n11/full/nn.2202.html.

Zhang, N. R. and von der Heydt, R. (2010). "Analysis of the context integration mechanisms underlying figure-ground organization in the visual cortex". In: *The Journal of Neuroscience* 30.19, pp. 6482–6496.

Zhaoping, L. (2005). "Border ownership from intracortical interactions in visual area V2". In: *Neuron* 47.1, pp. 143–153.

Zhou, H., Friedman, H. S., and von der Heydt, R. (2000). "Coding of border ownership in monkey visual cortex". In: *The Journal of Neuroscience* 20.17, pp. 6594–6611.

Vita

Jonathan R. Williford

Ph.D. Candidate

Johns Hopkins University, School of Medicine

Department of Neuroscience

Office: 338 Krieger Hall; 3400 N. Charles St.; Baltimore, MD, 21218-2685

Email: williford@jhu.edu

Thesis

Title: Neural basis of perceptual organization of natural scenes: Emergence of object-based coding in the primate visual cortex.

Advisor: Rüdiger von der Heydt

Committee: Marshall G. Hussain Shuler, Ernst Niebur, Stewart H.C. Hendry, and Howard Egeth

VITA

Education

Ph.D., Neuroscience 8/2009 - 11/2014 (Expected)

Johns Hopkins University, School of Medicine, Baltimore, MD

M.S., Computer Science, Machine Learning Track 1/2006 - 8/2009

Columbia University, New York, NY

B.S., Computer Science with Honors 9/2001 - 12/2004

Purdue University, West Lafayette, IN

Grants and Awards

- VSS Travel Grant ($500) 5/2014
- Best Poster in Systems Neuroscience, Department Retreat ($100) 9/2012
- Visual Neuroscience Training Grant ($77,906) 1/2010 - 12/2010
- Travel for Okinawa Computational Neuroscience Course 6/2010
- Fred and Ruth Kort Young Scholar Award ($1,500) 12/2009

VITA

Publications

Peer-reviewed

- **Jonathan Williford** and Rüdiger von der Heydt. (2013) Border-ownership coding. Scholarpedia, 8(10), 30040. DOI: 10.4249/scholarpedia.30040.

Proceedings

- **Jonathan Williford**, Chintan Dalal, and Minbo Shim (2009). Spatial multimodal mean background model for real-time MTI. Proceedings SPIE 7338, Acquisition, Tracking, Pointing, and Laser Systems Technologies XXIII, 73380A. DOI: 10.1117/12.818969.

Presentations

Oral presentations

- **Jonathan Williford** and Rüdiger von der Heydt (2014, November). Early visual cortex consistently estimates border-ownership in simple figures and natural scenes. Society for Neuroscience, Washington, D.C., USA.

- **Jonathan Williford** and Rüdiger von der Heydt (2014, May). Early visual

cortex assigns border ownership in natural scenes according to image context. Visual Sciences Society, St. Pete Beach, Florida, USA.

- **Jonathan Williford** and Rüdiger von der Heydt (2013, August). Early visual cortex assigns border ownership in natural scenes according to image context. European Conference on Visual Perception, Bremen, Germany. (Invited)

- **Jonathan Williford** and Rüdiger von der Heydt (2012, May). Neural coding of border-ownership in natural scenes. Visual Science Society, Naples, FL, USA.

- **Jonathan Williford**, Chintan Dalal, and Minbo Shim (2009, April). Spatial multimodal mean background model for real-time MTI. In S.L Chodos & W.E. Thompson (Chairs) Target Detection and Tracking. Symposium conducted at SPIE Defense, Security, and Sensing Conference, Orlando, FL, USA.

Poster Presentations

- **Jonathan Williford** and Rüdiger von der Heydt (2012). Neural border ownership coding by the global context of natural scenes. Poster No. 464.15/Z8 New Orleans, LA: Society for Neuroscience.

- **Jonathan Williford**, Hee-Kyoung Ko, and Rüdiger von der Heydt (2011). Border ownership coding of natural scenes. Poster No. 271.09/II28 Washington, DC: Society for Neuroscience.

VITA

Work Experience

Senior Software Engineer, GD Robotic Systems **9/2006 - 8/2009**

Researched and solved problems using computer vision and machine learning for autonomous robots.

- Designed and implemented multiple moving target indication (MTI) systems, including one that used structure from motion (SFM) to calculate the speed of moving targets from a moving platform.
- Implemented algorithms that helped with MTI, such as MeanShift for image segmentation, multimodal mean background model by Apewokin et al., ensemble tracking by Avidan, shaped-based tracking, and Bayesian terrain classifier.
- Implemented a pedestrian detection system that fuses LADAR depth data with shape and color information from a camera.
- Implemented a generic simulated annealing framework in C++ to tune parameters of various modules, including background model and MTI.

Software Developer / Project Lead, gh LLC **12/2004 - 8/2006**

Led the MathSpeak project to automatically render mathematics aurally to visually impaired persons, which received a multimillion dollar grant from the Indiana 21st Century Research and Technology Fund.

- Defined the MathSpeak specification in collaboration with Dave Schleppenbach, the co-founder of gh LLC, and Dr. Abraham Nemeth, the inventor of the Nemeth Code, the Braille system for writing mathematics that is used in the United States and several other countries.

- Implemented software to automatically render equations written in MathML (an XML markup for mathematic equations) into text.

- Guided Cepstral LLC in developing of custom voices for MathSpeak.

- Worked with Dr. Lyle Lloyd, Mick Isaacson, and Samuel Mathew to design an experiment to study of efficacy of MathSpeak, which showed that MathSpeak was significantly less ambiguous than regularly spoken mathematics. Lead to multiple publications.

- As part of the MathML Modular Extension Working Group, developed the DAISY 3 Modular Extension for MathML which was approved on February 23, 2007. This extension provided a way for including mathematic equations into the DAISY/NISO Digital-Talking-Book specification, an XML specification for visually impaired persons.

Teaching Experience

Teaching assistant for Visual System (AS.080.355) 8/2011 - 12/2011

VITA

Johns Hopkins University

Professional Memberships

Society for Neuroscience Student Member	2011 - present
Vision Sciences Society Predoctoral Member	2012 - present
IEEE Student Member	9/2007 - 2/2009

Related Experience

Associate Editor at Scholarpedia	12/2013 - present
Assistant Editor at Scholarpedia	7/2010 - 12/2013

Skills

Electrophysiology, R, C/C++, Python, MATLAB, OpenCV Library, Linux, Windows, Portable Batch System (PBS), Git, Subversion, LaTeX